

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 10:29:22 : Search time 168.5 seconds  
(without alignments)  
1912.600 Million cell updates/sec

Title: US-09-980-265-9  
Perfect score: 20  
Sequence: 1 ttagtcttagttagttagt 20

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 60155 seqs, 8056825618 residues

Total number of hits satisfying chosen parameters: 120310

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: GenEmbl\_HTG:\*

1: gb\_htg:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	95.0	186542	1	AL844197 Danio rer
2	18.4	92.0	5421	1	AC019958 Drosophi
3	18.4	92.0	38069	1	AC017255 Drosophi
4	18.4	92.0	80679	1	AC103576 Rattus no
5	18.4	92.0	106651	1	AC106466 Rattus no
6	18.4	92.0	155877	1	AC128449 Rattus no
7	18.4	92.0	176400	1	AC113369 Homo sapi
8	18.4	92.0	178970	1	AC012589 Homo sapi
9	18.4	92.0	187431	1	AC023241 Homo sapi
10	18.4	92.0	213308	1	AC079164 Mus muscu
11	18.4	92.0	217953	1	AC122334 Mus muscu
12	18.4	92.0	241585	1	AL844536 Mus muscu
13	17.4	87.0	15997	1	AC015120 Drosophi
14	17.4	87.0	102832	1	AL713963 Danio rer
15	17.4	87.0	147734	1	AL713862 Danio rer
16	17.4	87.0	148208	1	AC010951 Homo sapi
17	17.4	87.0	153328	1	AC122111 Rattus no
18	17.4	87.0	159295	1	CNS079P3 Oryza sat
19	17.4	87.0	167683	1	AC079251 Homo sapi
20	17.4	87.0	169179	1	AC111888 Rattus no
21	17.4	87.0	172006	1	AC110979 Rattus no
22	17.4	87.0	173471	1	AC101841 Mus muscu
23	17.4	87.0	183450	1	AC119322 Rattus no
24	17.4	87.0	186581	1	AC130415 Homo sapi
25	17.4	87.0	195075	1	AC122291 Mus muscu
26	17.4	87.0	195444	1	AC107399 Homo sapi
27	17.4	87.0	209572	1	AC099615 Mus muscu
28	17.4	87.0	219491	1	AC092256 Mus muscu
29	17.4	87.0	225410	1	AC113592 Mus muscu
30	17.4	87.0	227884	1	AC114995 Mus muscu
31	17.4	87.0	247102	1	AC123854 Mus muscu
32	17.4	87.0	263211	1	AC098458 Rattus no

C 33	17.4	87.0	263546	1	AC099416 Mus muscu
C 34	17.4	87.0	265861	1	AC069274 Mus muscu
C 35	17.4	87.0	267052	1	AC122872 Mus muscu
C 36	17.4	87.0	276082	1	AC112365 Rattus no
C 37	16.8	84.0	8659	1	AC117270 Dictyoste
C 38	16.8	84.0	14921	1	AC015356 Drosophi
C 39	16.8	84.0	15707	1	AC019964 Drosophi
C 40	16.8	84.0	15792	1	AC014334 Drosophi
C 41	16.8	84.0	35974	1	AC015005 Drosophi
C 42	16.8	84.0	44765	1	AC116467 Mus muscu
C 43	16.8	84.0	50100	1	AC101025 Mus muscu
C 44	16.8	84.0	52486	1	AC101922 Mus muscu
C 45	16.8	84.0	55036	1	AC100584 Mus muscu

## ALIGNMENTS

RESULT 1  
LOCUS AL844197/c  
DEFINITION Danio rerio clone DKEX-156L9, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 7  
unorderd pieces.  
ACCESSION AL844197.2 GI:21953176  
VERSION HTG: HTGS\_PHASE1, HTGS\_DRAFT, HTGS\_FULLTOP.  
KEYWORDS  
SOURCE Danio rerio.  
ORGANISM Danio rerio.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.  
1 (bases 1 to 186542)  
Burton, J.  
Direct Submission  
Submitted (18-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zface@sanger.ac.uk  
Clone requests: clonerequest@sanger.ac.uk  
On Jul 24, 2002 this sequence version replaced gi:21912371.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: zface@sanger.ac.uk  
----- Project Information  
Center project name: zki56L9  
----- Summary Statistics  
Assembly program: XGAP4; Version 4.5  
Chemistry: Dye-terminator; 100% of reads  
Consensus quality: 184125 bases at least Q40  
Consensus quality: 184667 bases at least Q30  
Consensus quality: 185376 bases at least Q20  
Insert size: 185942; sum-of-ctrls  
Insert size: 189146; 2.5% error; agarose-1p  
Quality coverage: 7.02x in Q20 bases; sum-of-ctrls quality  
coverage: 6.90x in Q20 bases; agarose-1p  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 7 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 82061: contig of 82061 bp in length  
\* 82062 82161: gap of 100 bp  
\* 82162 92135: contig of 9974 bp in length  
\* 92136 92235: gap of 100 bp  
\* 92236 123057: contig of 30822 bp in length  
\* 123058 123157: gap of 100 bp  
\* 123158 133138: contig of 9961 bp in length  
\* 133139 133238: gap of 100 bp  
\* 133239 138160: contig of 4922 bp in length

FEATURES

source

1. 186542

/organism="Dario rerio"

/db\_xref="taxon:7955"

/clone\_lib="DKEY-156L9"

/clone\_lib="Dantiokey"

1. 82061

/note="assembly\_fragment:00649"

fragment\_chain:1"

82162. 92135

/note="assembly\_fragment:00405"

fragment\_chain:1"

92236. 123057

/note="assembly\_fragment:01992"

fragment\_chain:1"

123158. 133138

/note="assembly\_fragment:01602"

fragment\_chain:1"

133239. 138160

/note="assembly\_fragment:00752"

fragment\_chain:1"

138261. 175279

/note="assembly\_fragment:00271"

fragment\_chain:1"

175380. 186542

/note="assembly\_fragment:02718"

fragment\_chain:1"

BASE COUNT 61324 a 34425 c 33413 g 56780 t 600 others

ORIGIN

Query Match 95.0%; Score 19; DB 1; Length 186542;

Best Local Similarity 100.0%; Pred. No. 7.8;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TAGTCTTACTTATTAGTT 20

Db 20975 TAGTCTTACTTATTAGTT 20957

RESULT 2

AC019958/c 5421 bp DNA linear HTG 03-JAN-2000

LOCUS Drosophila melanogaster, \*\*\* SEQUENCING IN PROGRESS \*\*\*

DEFINITION pieces.

AC019958

AC019958.1 GI:6664939

VERSION HTG; HTGS\_PHASE2.

KEYWORDS Drosophila melanogaster.

SOURCE Drosophila melanogaster.

ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 5421)

AUTHORS Adams,M. and Venter,J.C.

TITLE Direct Submission

JOURNAL Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA

COMMENT This sequence was identified as CDM:10211530 by the submitter. For more information on this record e-mail to fly@celera.com.

\* NOTE: This is a 'working draft' sequence.

\* This sequence will be replaced

\* by the finished sequence as soon as it is available and

\* the accession number will be preserved.

FEATURES

source

1. 5421

/organism="Drosophila melanogaster"

/db\_xref="taxon:7227"

BASE COUNT 1558 a 1150 c 1146 g 1567 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 5421;

Best Local Similarity 95.0%; Pred. No. 14;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTACTTATTAGTT 20

Db 435 TTAGTCTTACTTATTAGTT 416

RESULT 3

AC017255 38069 bp DNA linear HTG 09-DEC-1999

LOCUS Drosophila melanogaster, \*\*\* SEQUENCING IN PROGRESS \*\*\*

DEFINITION pieces.

AC017255

AC017255.1 GI:6553731

VERSION HTG; HTGS\_PHASE2.

KEYWORDS Drosophila melanogaster.

SOURCE Drosophila melanogaster.

ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 38069)

AUTHORS Adams,M. and Venter,J.C.

TITLE Direct Submission

JOURNAL Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA

COMMENT This sequence was identified as CDM:10210065 by the submitter. For more information on this record e-mail to fly@celera.com.

\* NOTE: This is a 'working draft' sequence.

\* This sequence will be replaced

\* by the finished sequence as soon as it is available and

\* the accession number will be preserved.

FEATURES

source

1. 38069

/organism="Drosophila melanogaster"

/db\_xref="taxon:7227"

BASE COUNT 12432 a 7322 c 7198 g 11117 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 38069;

Best Local Similarity 95.0%; Pred. No. 16;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTACTTATTAGTT 20

Db 2372 TTAGTCTTACTTATTAGTT 2391

RESULT 4

AC103576 80679 bp DNA linear HTG 12-JUL-2002

LOCUS Rattus norvegicus clone CH230-63N13, \*\*\* SEQUENCING IN PROGRESS

DEFINITION \*\*\*47 unordered pieces.

AC103576

AC103576.4 GI:21729709

VERSION HTG; HTGS\_PHASE1.

KEYWORDS Norway rat.

SOURCE Rattus norvegicus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 80679)

AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., All-ouman,F.R., Allen,C., Alsbrooks,S.L., Amaralunga,H.C., Are,J.R., Ayele,M., Banks,T., Barbarta,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buhaq,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,

Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Eamhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
Homs,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,U., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.,  
Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,  
Locado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheswari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M.,  
Miner,G., Miner,Z., Mitchell,T., Monabhat,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokwuo,S., Ogun,M., Okunolu,G.,  
Oregunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pu,L., Quiles,S., Ren,Y.,  
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,  
Schreier,S., Scott,G., Shen,H., Shooshitari,N., Sisson,I.,  
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Swalek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Usmani,K., Vasquez,L., Vera,Y., Villalón,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,  
Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G. and Gibbs,R.

Unpublished  
2 (bases 1 to 80679)  
Worley,K.C.

Direct Submission  
Submitted (29-NOV-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 80679)  
Worley,K.C.

Direct Submission  
Submitted (12-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 11, 2002 this sequence version replaced gi:18846139.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: [http://www.hgsc.bcm.tmc.edu/docs/genbank/draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html)/  
Contact: hgsc-help@bcm.tmc.edu

----- Project Information  
Center project name: GIGY  
Center clone name: CH230-63N13

----- Summary Statistics  
Chemistry: Dye-terminator Big Dye: 100% of reads  
Assembly program: Phrap: version 0.990329  
Consensus quality: 33297 bases at least Q40  
Consensus quality: 35946 bases at least Q30  
Consensus quality: 38652 bases at least Q20

-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank/draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 47 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1171: contig of 1171 bp in length  
1172 1271: gap of unknown length

1272 2868: contig of 1597 bp in length  
2869 4131: contig of unknown length  
4132 4231: contig of 1163 bp in length  
4232 5312: contig of 1081 bp in length  
5313 5412: contig of unknown length  
5413 6870: contig of 1458 bp in length  
6871 7991: contig of unknown length  
7992 8091: contig of 1021 bp in length  
8092 9126: contig of unknown length  
9127 9227: contig of 1035 bp in length  
9228 10468: contig of 1242 bp in length  
10469 10569: contig of unknown length  
10570 11620: contig of 1051 bp in length  
11621 11719: contig of unknown length  
11720 12894: contig of 1175 bp in length  
12895 12994: contig of unknown length  
12995 14178: contig of 1184 bp in length  
14179 14278: contig of unknown length  
14279 15615: contig of 1337 bp in length  
15616 15715: contig of unknown length  
15716 17016: contig of 1301 bp in length  
17017 17116: contig of unknown length  
17118 18233: contig of 1117 bp in length  
18234 18333: contig of unknown length  
18334 19665: contig of 1332 bp in length  
19666 19765: contig of unknown length  
19766 21151: contig of 1386 bp in length  
21152 22739: contig of unknown length  
22740 22839: contig of 1488 bp in length  
22840 23909: contig of 1070 bp in length  
23910 24009: contig of unknown length  
24010 25502: contig of 1493 bp in length  
25503 25602: contig of unknown length  
25603 26796: contig of 1194 bp in length  
26797 26896: contig of unknown length  
26897 28120: contig of 1224 bp in length  
28121 28220: contig of unknown length  
28221 29944: contig of 1724 bp in length  
29945 30044: contig of unknown length  
30045 32238: contig of 2194 bp in length  
32239 32338: contig of unknown length  
32340 33569: contig of 1231 bp in length  
33570 33669: contig of unknown length  
33670 35660: contig of 1891 bp in length  
35661 35660: contig of unknown length  
35662 36788: contig of 1128 bp in length  
36789 36888: contig of unknown length  
36889 38239: contig of 1351 bp in length  
38240 38339: contig of unknown length  
38340 39574: contig of 1235 bp in length  
39575 39674: contig of unknown length  
39676 40719: contig of 1045 bp in length  
40720 42444: contig of 1625 bp in length  
42445 42544: contig of unknown length  
42545 44516: contig of 1972 bp in length  
44517 44616: contig of unknown length  
44618 46392: contig of 1776 bp in length  
46393 46492: contig of unknown length  
46493 47848: contig of 1356 bp in length  
47849 47948: contig of unknown length  
47949 49380: contig of 1432 bp in length  
49381 49480: contig of unknown length  
49481 50992: contig of 1512 bp in length  
50993 53214: contig of 2122 bp in length  
53215 53314: contig of unknown length  
53316 54686: contig of 1372 bp in length  
54687 54786: contig of unknown length  
54787 56476: contig of 1690 bp in length

FEATURES			
*	56477	56576:	gap of unknown length
*	56577	56624:	contig of 2048 bp in length
*	56625	56724:	gap of unknown length
*	56725	61311:	contig of 2587 bp in length
*	61312	61411:	gap of unknown length
*	61412	63407:	contig of 1996 bp in length
*	63408	63507:	gap of unknown length
*	63508	65412:	contig of 1905 bp in length
*	65413	65512:	gap of unknown length
*	65513	68121:	contig of 2609 bp in length
*	68122	68221:	gap of unknown length
*	68222	71194:	contig of 2973 bp in length
*	71195	71294:	gap of unknown length
*	71295	73654:	contig of 2360 bp in length
*	73655	73754:	gap of unknown length
*	73755	76459:	contig of 2705 bp in length
*	76460	76559:	gap of unknown length
*	76560	80679:	contig of 4120 bp in length
		Location/Qualifiers	

BASE COUNT	17902	a	18300	c	17407	g	18485	t	8585	others
ORIGIN										

Query Match	92.0%	Score 18.4	DB 1	Length 80679
Best Local Similarity	95.0%	Pred No. 17		
Matches 19, Conservative	0	Mismatches 1	Indels 0	Gaps 0

QY	1	TTAGTCTCTAGTTATTAGTT	20
Db	41343	TTAGTTATTAGTTATTAGTT	41362

RESULT 5	
AC106466/c	
LOCUS	AC106466 106651 bp DNA linear HTG 13-JUL-2007
DEFINITION	Rattus norvegicus clone CHZ30-21B18, *** SEQUENCING IN PROGRESS
	**, 45 unordered pieces.
ACCESSION	AC106466
VERSION	AC106466.2 GI:21735240
KEYWORDS	HTG; HTGS_PHASE1.
SOURCE	Norway rat.
ORGANISM	Rattus norvegicus

## REFERENCE

1 (bases 1 to 106651)

Muzny, D. M., Adams, C., Adio-Oduola, B., Ali-osman, F. R., Allen, C., Alstbrooks, S. L., Amaral-Lungue, H. C., Are, J. R., Ayela, M., Banks, T., Barbieri, J., Benton, J., Bimoge, K., Blankenhuy, K., Bonnah, D., Bouch, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N. P., Buhey, C., Burch, P., Burkett, C., Burrell, K. L., Byrd, N. C., Carrott, T. F., Carter, M., Cavazos, S. R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C. D., Cox, C., Coyle, M. D., Dathorne, S. R., David, R., Davila, M. L., Davis, C., Davy-Carroll, L., Deedrich, D. A., Delaney, K. R., Delgado, O., Demm, A. L., Ding, Y., Dinh, H. J., Douthett, K. J., Draper, H., Dugan-Rocha, S., Durbin, K. J., Earnhart, C., Edgar, D., Edwards, C. C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J. H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, A., Hernandez, O., Hernandez, O., Hodgson, A., Hognues, M., Holloway, C., Hollins, B., Homsi, F., Howard, S., Huber, J., Huily, S., Hume, J., Jackson, L. E., Jacobson, B., Jia, T., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelli, S., Khan, U., King, L., Koryah, J., Kover, C., Kratovic, J., Kureshi, A., Landry, N., Leel, B., Lewis, L. C., Lewis, L. J., Li, Z., Licharge, O., Liu, C., Liu, J., Liu, W., Louisgeed, H., Lozdo, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E.,

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

COMMENT

Massey, E., Mawhinney, E., McLeod, M. P., Meador, M., Mei, G., Melker, M.,  
 Moser, M., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,  
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
 Nguyen, N., Nickerson, E., Nwokwenkwo, S., Ogun, M., Okwuonu, G.,  
 Oragunye, N., Oviyedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
 Peters, L., Pickens, R., Primus, E., Pu, L. L., Quiles, M., Ren, Y.,  
 Rivers, M., Rojas, A., Rotjoubkan, I., Rolfe, M., Ruiz, S., Savery, G.,  
 Scheraga, S., Scott, G., Shen, H., Shoochait, N., Sisson, I.,  
 Sodergren, E., Sonake, T., Sparks, A., Stanley, H., Stone, H.,  
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
 Tansey, J., Taylor, C., Taylor, T., Teifrod, B., Thomas, R., Thomas, S.,  
 Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,  
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
 Williams, G., Williamson, A., Wleciyk, R., Wooden, S., Worley, K.,  
 Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,  
 Weinstein, G., and Gibbs, R.  
 -----  
 Direct Submission  
 2 (bases 1 to 106651)  
 Unpublished  
 Worley, K.C.  
 Direct Submission  
 Submitted (12-JAN-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 106651)  
 Worley, K.C.  
 -----  
 Direct Submission  
 Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Jul 12, 2002 this sequence version replaced gi:18138988.  
 -----  
 Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 -----  
 Project Information  
 Center project name: GKyr  
 Center clone name: CH230-211B18  
 -----  
 Summary Statistics  
 Sequencing vector: Plasmid:  
 Chemistry: Dye-terminator Big Dye: 100% of reads  
 Assembly program: Phrap: version 0.990329  
 Consensus quality: 79275 bases at least Q40  
 Consensus quality: 83497 bases at least Q30  
 Consensus quality: 87549 bases at least Q20  
 -----  
 NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/genbank/draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html))  
 NOTE: This is a 'working draft' sequence. It currently  
 consists of 45 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.  
 -----  
 1 1130: contig of 1130 bp in length  
 1131 1230: gap of unknown length  
 1231 1243: contig of 1013 bp in length  
 2244 2343: gap of unknown length  
 2344 3368: contig of 1026 bp in length  
 3370 3469: gap of unknown length  
 3470 5003: contig of 1534 bp in length  
 5004 5103: gap of unknown length  
 5104 6657: contig of 1554 bp in length  
 6658 6757: gap of unknown length  
 6758 7840: contig of 1083 bp in length  
 7841 7940: gap of unknown length  
 7941 10020: contig of 2080 bp in length  
 10021 10120: gap of unknown length  
 10121 11477: contig of 1357 bp in length  
 11478 11577: gap of unknown length



11578 13365: contig of 1788 bp in length  
13366 13465: gap of unknown length  
13466 14853: contig of 1388 bp in length  
14854 14953: gap of unknown length  
14954 16290: contig of 1337 bp in length  
16291 16390: gap of unknown length  
16391 17509: contig of 1119 bp in length  
17510 17609: gap of unknown length  
17610 19000: contig of 1391 bp in length  
19001 19100: gap of unknown length  
19101 20944: contig of 1844 bp in length  
20945 21044: gap of unknown length  
21045 22651: contig of 1607 bp in length  
22652 22751: gap of unknown length  
22752 25058: contig of 2307 bp in length  
25059 25158: gap of unknown length  
25159 26816: contig of 1658 bp in length  
26817 26916: gap of unknown length  
26917 28641: contig of 1725 bp in length  
28642 28741: gap of unknown length  
28742 30799: contig of 2058 bp in length  
30800 30899: gap of unknown length  
30900 32351: contig of 1452 bp in length  
32352 32451: gap of unknown length  
32452 34522: contig of 2071 bp in length  
34523 34622: gap of unknown length  
34623 36509: contig of 1887 bp in length  
36510 36609: gap of unknown length  
36610 38251: contig of 1642 bp in length  
38252 38351: gap of unknown length  
38352 40309: contig of 1958 bp in length  
40310 40409: gap of unknown length  
40410 42792: contig of 2383 bp in length  
42793 42892: gap of unknown length  
42893 45780: contig of 2888 bp in length  
45781 45881: gap of unknown length  
45882 49221: contig of 3341 bp in length  
49222 49321: gap of unknown length  
49323 51468: contig of 2147 bp in length  
51469 51568: gap of unknown length  
51569 53749: contig of 2181 bp in length  
53750 53849: gap of unknown length  
53850 55748: gap of unknown length  
55749 55848: contig of 1899 bp in length  
55849 59788: contig of 3940 bp in length  
59789 59888: gap of unknown length  
59889 62649: contig of 2761 bp in length  
62650 62749: gap of unknown length  
62750 65930: contig of 3181 bp in length  
65931 66030: gap of unknown length  
66031 68484: contig of 2454 bp in length  
68485 68584: gap of unknown length  
68585 70509: contig of 1925 bp in length  
70510 70609: gap of unknown length  
70610 73606: contig of 2997 bp in length  
73607 73706: gap of unknown length  
73707 76448: contig of 2742 bp in length  
76449 76548: gap of unknown length  
76550 79392: contig of 2844 bp in length  
79393 79492: gap of unknown length  
79493 83319: contig of 2827 bp in length  
83320 82419: gap of unknown length  
82420 85792: contig of 3373 bp in length  
85793 85892: gap of unknown length  
85893 88433: contig of 2541 bp in length  
88434 88533: gap of unknown length  
88534 92245: contig of 3712 bp in length  
92246 92345: gap of unknown length  
92346 96997: contig of 4652 bp in length  
96998 97097: gap of unknown length  
97098 102010: contig of 4913 bp in length  
102011 102110: gap of unknown length  
102111 106651: contig of 4541 bp in length.

FEATURES  
source Location/Qualifiers  
1..106651  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/clone="CH230-211B18"

BASE COUNT 29303 a 19724 c 19632 g 30261 t 7731 others  
ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 106651;  
Best Local Similarity 95.0%; Pred. No. 17;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTATTACTT 20  
||||| |||||||||  
DB 83573 TTAGTATTAGTATTACTT 83554

RESULT 6  
AC128449  
LOCUS  
DEFINITION Rattus norvegicus clone CH230-117D11, \*\*\* SEQUENCING IN PROGRESS  
AC128449 155877 bp DNA linear HTG 19-JUL-2002  
AC128449.1 GI:21909143  
VERSION  
KEYWORDS HTG: HTGS\_PHASE1.  
SOURCE Rattus norvegicus  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 155877)  
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
Albrooks,S.L., Amaralunge,H.C., Are,J.R., Ayete,M., Banks,T.,  
Barbata,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,  
Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,D., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Devila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.,  
Douthwaite,K.J., Draper,H., Dugan-Kocha,S., Durbin,K.J.,  
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorrell,J.H., Guevara,W., Gunaratne,P., Hate,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hoggues,M., Holloway,C., Hollins,B.,  
Honsl,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlssoen,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,M., Louisged,H.,  
Lorado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokkwo,S., Ogun,M., Okunnu,G.,  
Oreagunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Puy,L.L., Quiles,M., Ren,Y.,  
Rivers,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savary,G.,  
Scherer,S., Scott,G., Shen,H., Shooshitari,N., Sisson,I.,  
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Swatek,A., Tabor,P., Tameris,A., Tameisa,K., Tang,H.,  
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
Williams,G., Williamson,A., Wleciak,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorilla,S., Nelson,D.,  
Weinstock,G. and Gibbs,R.

TITLE  
JOURNAL  
Unpublished

REFERENCE 2 (bases 1 to 155877)  
 AUTHORS Worley, K.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GWL  
 Center clone name: CH230-117D11  
 ----- Summary Statistics  
 Sequencing vector: plasmid;  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 73163 bases at least Q40  
 Consensus quality: 82138 bases at least Q30  
 Consensus quality: 87186 bases at least Q20  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 85 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1063	1062: contig of 1062 bp in length	28405	28504: gap of unknown length
1163	1162: gap of unknown length	28505	29670: contig of 1166 bp in length
2171	2171: contig of 1009 bp in length	29671	29770: gap of unknown length
2272	2271: gap of unknown length	29771	30994: contig of 1224 bp in length
3322	3321: contig of 1050 bp in length	30995	31094: gap of unknown length
3421	3421: gap of unknown length	31095	32675: contig of 1561 bp in length
4463	4463: contig of 1042 bp in length	32676	32775: gap of unknown length
4563	4563: gap of unknown length	32776	34130: contig of 1355 bp in length
4564	4563: contig of 1024 bp in length	34131	34230: gap of unknown length
5587	5587: gap of unknown length	34231	35489: contig of 1259 bp in length
5687	5687: gap of unknown length	35490	35589: gap of unknown length
6712	6712: contig of 1025 bp in length	35590	37026: contig of 1437 bp in length
6713	6812: gap of unknown length	37027	37126: gap of unknown length
6813	7822: contig of 1010 bp in length	37127	38796: contig of 1670 bp in length
7823	7822: gap of unknown length	38797	38896: gap of unknown length
7923	9753: contig of 1831 bp in length	38897	40001: contig of 1105 bp in length
9754	9853: gap of unknown length	40002	40101: gap of unknown length
9854	10887: contig of 1034 bp in length	40102	41115: contig of 1014 bp in length
10888	10987: gap of unknown length	41116	41215: gap of unknown length
10988	12401: contig of 1414 bp in length	41216	42569: contig of 1354 bp in length
12402	12501: gap of unknown length	42570	42669: gap of unknown length
12502	14037: contig of 1536 bp in length	42670	43830: contig of 1161 bp in length
14038	14137: gap of unknown length	43831	43930: gap of unknown length
14138	15197: contig of 1060 bp in length	43931	45281: contig of 1351 bp in length
15198	15297: gap of unknown length	45282	45381: gap of unknown length
15298	16644: contig of 1347 bp in length	45382	47441: contig of 2060 bp in length
16645	16744: gap of unknown length	47442	47541: gap of unknown length
16745	18177: contig of 1433 bp in length	47542	49026: contig of 1485 bp in length
18178	18277: gap of unknown length	49027	49126: gap of unknown length
18278	19658: contig of 1381 bp in length	49127	50263: contig of 1137 bp in length
19659	19758: gap of unknown length	50264	50363: gap of unknown length
19759	20806: contig of 1048 bp in length	50364	52441: contig of 2078 bp in length
20807	20906: gap of unknown length	52442	52541: gap of unknown length
20907	22312: contig of 1406 bp in length	52542	53981: contig of 1440 bp in length
22313	22412: gap of unknown length	53982	54081: gap of unknown length
22413	23922: contig of 1510 bp in length	54082	55502: contig of 1421 bp in length
23923	24022: gap of unknown length	55503	55602: gap of unknown length
24023	25433: contig of 1411 bp in length	55603	56893: contig of 1251 bp in length
25434	25533: gap of unknown length	56894	56993: gap of unknown length
25534	26809: contig of 1276 bp in length	56994	58953: contig of 1960 bp in length
26810	26909: gap of unknown length	58954	59053: gap of unknown length
26910	28404: contig of 1495 bp in length	59054	60351: contig of 1298 bp in length
		60352	60451: gap of unknown length
		60452	61848: contig of 1397 bp in length
		61849	61948: gap of unknown length
		61949	64132: contig of 2184 bp in length
		64133	64232: gap of unknown length
		64233	65624: contig of 1392 bp in length
		65625	65724: gap of unknown length
		65725	66795: contig of 1071 bp in length
		66796	66895: gap of unknown length
		66896	68161: contig of 1266 bp in length
		68162	68261: gap of unknown length
		68262	69390: contig of 1129 bp in length
		69391	69490: gap of unknown length
		69491	70785: contig of 1295 bp in length
		70786	70885: gap of unknown length
		70886	73269: contig of 2384 bp in length
		73270	73369: gap of unknown length
		73370	75050: contig of 1681 bp in length
		75051	75150: gap of unknown length
		75151	76285: contig of 1135 bp in length
		76286	76385: gap of unknown length
		76386	78166: contig of 1781 bp in length
		78167	78266: gap of unknown length
		78267	80185: contig of 1919 bp in length
		80186	80285: gap of unknown length
		80286	81681: contig of 1396 bp in length
		81682	81781: gap of unknown length
		81782	83340: contig of 1559 bp in length
		83341	83440: gap of unknown length
		83441	85436: contig of 1996 bp in length
		85437	85536: gap of unknown length

Query Match 92.0%; Score 18.4; DB 1; Length 155877;  
Best Local Similarity 95.0%; Pred. No. 18;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTTATTACTT 20  
|||||  
Db 46393 TTAGTCTTACTTACTTACTT 46412

RESULT 7  
AC113369 176400 bp DNA linear HTG 01-MAR-2002  
LOCUS  
DEFINITION Homo sapiens chromosome 5 clone RP11-119P10, WORKING DRAFT  
SEQUENCE, 1 unordered piece.  
AC113369  
AC113369.1 GI:19033483  
VERSION  
KEYWORDS HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ACTIVEFIN.  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 176400)  
DOE Joint Genome Institute.  
Sequencing of Human Chromosome 5  
Unpublished  
2 (bases 1 to 176400)  
DOE Joint Genome Institute.  
Direct Submission  
Submitted (01-MAR-2002) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
-----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: http://www.jgi.doe.gov

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Project Information  
Center Project Name: 452431  
Center clone name: RPCI-11\_119P10  
-----

Summary Statistics  
Consensus quality: 175634 bases at least Q40  
Consensus quality: 176221 bases at least Q30  
Consensus quality: 176375 bases at least Q20  
Estimated insert size: 182000; agarose-fp estimation  
Estimated insert size: 176400; sum-of-contigs estimation  
Quality coverage: 12.49 in Q20 bases; agarose-fp estimation  
Quality coverage: 12.88 in Q20 bases; sum-of-contigs estimation.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 1 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 176400: contig of 176400 bp in length.  
Location/Qualifiers  
1..176400  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/clone="RP11-119P10"  
/clone\_lib="RPCI human BAC library 11"  
BASE COUNT 55190 a 37584 c 34956 g 48670 t

FEATURES  
SOURCE

Query Match 92.0%; Score 18.4; DB 1; Length 176400;  
Best Local Similarity 95.0%; Pred. No. 18;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTTATTACTT 20  
|||||

Db 78311 TTAGTATTACTTATTACTT 78330

RESULT 8  
AC012589 178970 bp DNA linear HTG 24-AUG-2002  
LOCUS  
DEFINITION Homo sapiens chromosome 5 clone RP11-119P10 map 5, WORKING DRAFT  
SEQUENCE, 23 unordered pieces.  
AC012589  
AC012589.5 GI:7387335  
VERSION  
KEYWORDS HTG: HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 178970)  
Birken, B., Linton, L., Nusbaum, C. and Lander, E.  
Homo sapiens chromosome 5, clone RP11-119P10  
Unpublished  
2 (bases 1 to 178970)  
Birken, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,  
Baldwin, J., Barne, N., Beckert, R., Boguslavsky, L., Boukhalter, B.,  
Brown, A., Castle, A., Colangelo, M., Collins, S., Collamore, A.,  
Cooke, P., DeArrellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,  
Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,  
Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howard, J., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,  
Lehoczy, J., Lien, C., Locke, K., Macdonald, P., Margulis, N.,  
McEwan, P., McGurt, A., McKernan, K., McLaughlin, J., Meldrum, J.,  
Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Testaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,  
Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Direct Submission  
Submitted (30-OCT-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 178970)  
Birken, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,  
Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A.,  
Camarata, J., Campopiano, A., Chang, J., Choepel, Y., Colangelo, M.,  
Collins, S., Collamore, A., Cooke, P., DeArrellano, K., Dewar, K.,  
Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D.,  
Galagan, J., Gardyna, S., Glnde, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W.,  
Iliev, I., Johnson, R., Jones, C., Karatas, A., Laroque, K.,  
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,  
Maclean, C., Macdonald, P., Margulis, N., Matthews, C., McCarty, M.,  
McEwan, P., McKernan, K., McPheters, R., Meldrum, J., Menus, L.,  
Milnova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C.,  
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,  
Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,  
Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,  
Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnapp, R., Seaman, S.,  
Severy, P., Sounguez, C., Spencer, B., Stange-Thomann, N.,  
Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S.,  
Theodore, J., Travers, M., Travis, N., Trigglio, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,  
Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Apr 3, 2000 this sequence version replaced gi:6910762.  
All repeats were identified using RepeatMasker:  
http://ftp.genome.washington.edu/RW/RepeatMasker.html

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: MIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L3738

Center clone name: 119\_P-10  
----- Summary Statistics -----  
Sequencing vector: M13; M7815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 160436 bases at least Q40  
Consensus quality: 168973 bases at least Q30  
Consensus quality: 173202 bases at least Q20  
Insert size: 17700; agarose-fp  
Insert size: 176770; sum-of-ctrls  
Quality coverage: 4.0 in Q20 bases; agarose-fp  
Quality coverage: 4.0 in Q20 bases; sum-of-ctrls

NOTE: This is a 'working draft' sequence. It currently consists of 23 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1011: contig of 1011 bp in length  
1012 1111: gap of 100 bp  
1112 2722: contig of 1611 bp in length  
2723 2822: gap of 100 bp  
2823 4300: contig of 1478 bp in length  
4301 4400: gap of 100 bp  
4401 5425: contig of 1025 bp in length  
5426 5525: gap of 100 bp  
5526 7534: contig of 2009 bp in length  
7535 7634: gap of 100 bp  
7635 9780: contig of 2146 bp in length  
9781 9880: gap of 100 bp  
9881 12462: contig of 2582 bp in length  
12463 15626: gap of 100 bp  
15627 15726: gap of 100 bp in length  
15727 17635: contig of 1908 bp in length  
17636 17735: gap of 100 bp  
17736 19894: contig of 2159 bp in length  
19895 19994: gap of 100 bp  
19995 25298: contig of 5304 bp in length  
25299 25398: gap of 100 bp  
25399 30406: contig of 5008 bp in length  
30407 30506: gap of 100 bp  
30507 38666: contig of 8160 bp in length  
38667 38766: gap of 100 bp  
38767 45449: contig of 6683 bp in length  
45450 45549: gap of 100 bp  
45550 55180: contig of 9631 bp in length  
55181 55280: gap of 100 bp  
55281 63752: contig of 8472 bp in length  
63753 63852: gap of 100 bp  
63853 75295: contig of 11443 bp in length  
75296 75395: gap of 100 bp  
75396 87583: contig of 12188 bp in length  
87584 87683: gap of 100 bp  
87684 98245: contig of 10562 bp in length  
98246 98345: gap of 100 bp  
98346 111685: contig of 13340 bp in length  
111686 111785: gap of 100 bp  
111786 131002: contig of 19217 bp in length  
131003 131102: gap of 100 bp  
131103 152070: contig of 20968 bp in length  
152071 152170: gap of 100 bp  
152171 178970: contig of 26800 bp in length.

FEATURES  
Source

1.178970  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/map="5"  
/clone="RP11-119P10"

misc\_feature /clone.lib="RP11-11 Human Male BAC"  
1.1011  
/note="assembly-fragment"  
misc\_feature /note="assembly-fragment"  
1112.2722  
/note="assembly-fragment"  
misc\_feature 2823.4300  
/note="assembly-fragment"  
misc\_feature 4401.5425  
/note="assembly-fragment"  
misc\_feature 5526.7534  
/note="assembly-fragment"  
misc\_feature 7635.9780  
/note="assembly-fragment"  
misc\_feature 9881.12462  
/note="assembly-fragment"  
misc\_feature 12563.15626  
/note="assembly-fragment"  
misc\_feature 15727.17635  
/note="assembly-fragment"  
misc\_feature 17736.19894  
/note="assembly-fragment"  
misc\_feature 19995.25298  
/note="assembly-fragment"  
misc\_feature 25399.30406  
/note="assembly-fragment"  
misc\_feature 30507.38666  
/note="assembly-fragment"  
misc\_feature 38767.45449  
/note="assembly-fragment"  
misc\_feature 45550.55180  
/note="assembly-fragment"  
misc\_feature 55281.63752  
/note="assembly-fragment"  
misc\_feature 63853.75295  
/note="assembly-fragment"  
misc\_feature 75396.87583  
/note="assembly-fragment"  
misc\_feature 87684.98245  
/note="assembly-fragment"  
misc\_feature 98346.111685  
/note="assembly-fragment"  
misc\_feature 111786.131002  
/note="assembly-fragment"  
misc\_feature 131103.152070  
/note="assembly-fragment"  
misc\_feature /note="assembly-fragment"  
vector\_side:left"  
vector\_side:right"  
vector\_side:SP6  
clone\_end:77  
clone\_end:17  
misc\_feature /note="assembly-fragment"  
vector\_side:left"  
vector\_side:right"  
vector\_side:SP6  
clone\_end:77  
clone\_end:17  
misc\_feature /note="assembly-fragment"

BASE COUNT 54765 a 37234 c 35721 g 49050 t 2200 others

Query Match 92.0%; Score 18.4; DB 1; Length 178970;  
Best Local Similarity 95.0%; Pred. No. 18;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTATTAGTT 20  
||||| |||||||||  
Db 22509 TTAGTATTAGTATTAGTT 22528

RESULT 9  
AC023241/C 187431 bp DNA linear HTG 01-MAR-2000  
LOCUS Homo sapiens clone RP11-14L13, WORKING DRAFT SEQUENCE, 17 unordered  
DEFINITION pieces.  
ACCESSION AC023241  
VERSION AC023241.2 GI:7138122  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 187431)  
Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
Homo sapiens, clone RP11-14L13  
Unpublished  
2 (bases 1 to 187431)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,  
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,  
Chepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,  
DeArillano,K., Dewar,K., Domino,M., Doyle,M., Fennestor,J.,  
Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,  
Gardyna,S., Grant,G., Hagos,B., Heatford,A., Horton,L.,  
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,  
Lander,T., Lehoczy,J., Levine,R., Liu,C., Liu,G., Locke,K.,  
Mcdonald,P., Margulis,N., McEwan,P., McGurk,A., McKernan,K.,  
McPheters,R., Meldrim,J., Menius,L., Morrow,J., Naylor,J.,  
Norman,C.H., O'Connor,T., O'Donnell,P., Olivart,T.M., Peterson,K.,  
Pierre,N., Pisanl,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,  
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,  
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,  
Zimmer,A. and Zody,M.

TITLE  
JOURNAL  
COMMENT

Submitted (10-FEB-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Mar 1, 2000 this sequence version replaced gi:6958062.  
All repeats were identified using RepeatMasker:  
Smt, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research

Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu

Project Information  
Center project name: L3420  
Center clone name: 14.L.13

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 174185 bases at least Q40  
Consensus quality: 180351 bases at least Q30  
Consensus quality: 183357 bases at least Q20  
Insert size: 170000; agarose-ff  
Insert size: 185831; sum-of-ctnigs  
Quality coverage: 4.7 in Q20 bases; agarose-ff  
Quality coverage: 4.3 in Q20 bases; sum-of-ctnigs

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 17 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1126: contig of 1126 bp in length  
\* 1127 1226: gap of 100 bp  
\* 1227 2243: contig of 1017 bp in length  
\* 2244 2343: gap of 100 bp  
\* 2344 7309: contig of 4966 bp in length  
\* 7310 7409: gap of 100 bp  
\* 7410 11226: contig of 3817 bp in length  
\* 11227 11326: gap of 100 bp  
\* 11327 18500: contig of 7174 bp in length  
\* 18501 18600: gap of 100 bp  
\* 18601 27141: contig of 8541 bp in length  
\* 27142 27241: gap of 100 bp  
\* 27242 34289: contig of 7048 bp in length  
\* 34290 34389: gap of 100 bp

\* 34390 43865: contig of 9476 bp in length  
\* 43866 43965: gap of 100 bp  
\* 43966 53252: contig of 9287 bp in length  
\* 53253 53352: gap of 100 bp  
\* 53353 66152: contig of 12800 bp in length  
\* 66153 66252: gap of 100 bp  
\* 66253 73973: contig of 7721 bp in length  
\* 73974 74073: gap of 100 bp  
\* 74074 86391: contig of 12318 bp in length  
\* 86392 86491: gap of 100 bp  
\* 86492 98260: contig of 12769 bp in length  
\* 98261 99360: gap of 100 bp  
\* 99361 119628: contig of 20268 bp in length  
\* 119629 119728: gap of 100 bp  
\* 119729 142636: contig of 22908 bp in length  
\* 142637 142736: gap of 100 bp  
\* 142737 164452: contig of 21716 bp in length  
\* 164453 164552: gap of 100 bp  
\* 164553 187431: contig of 22879 bp in length.

## FEATURES

source

1. 187431  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="RP11-14L13"  
/clone\_1lb="RPC1-11 Human Male BAC"  
1. 1126  
/note="assembly-fragment"  
1227. 2243  
/note="assembly-fragment"  
2344. 7309  
/note="assembly-fragment"  
7410. 11226  
/note="assembly-fragment"  
11327. 18500  
/note="assembly-fragment"  
18601. 27141  
/note="assembly-fragment"  
27242. 34289  
/note="assembly-fragment"  
clone\_end:77  
vector\_side:right"  
34390. 43865  
/note="assembly-fragment"  
43966. 53252  
/note="assembly-fragment"  
53353. 66152  
/note="assembly-fragment"  
66253. 73973  
/note="assembly-fragment"  
74074. 86391  
/note="assembly-fragment"  
86492. 99260  
/note="assembly-fragment"  
clone\_end:96  
vector\_side:left"  
99361. 119628  
/note="assembly-fragment"  
119729. 142636  
/note="assembly-fragment"  
142737. 164452  
/note="assembly-fragment"  
164553. 187431  
/note="assembly-fragment"

## BASE COUNT

53296 a 39959 c 39617 g 52948 t 1611 others

## Query Match

Best Local Similarity 92.0%; Score 18.4; DB 1; Length 187431;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 TTAGTCTTACTTATTACTT 20

DB 106825 TTAGTCTTACTTATTACTT 106806

RESULT 10	AC079164	213308 bp	DNA	linear	HTG 22-Aug-2000
LOCUS	AC079164				
DEFINITION	Mus musculus chromosome 12 clone RP23-157F24 strain C57BL6/J.				
ACCESSION	AC079164				
VERSION	AC079164.1	GI:9864742			
KEYWORDS	HTG; HTGS; PHASE1; HTGS_DRAFT.				
SOURCE	Mus musculus.				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 213308)				
AUTHORS	Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Boulford,G.G., Dietrich,N.L., Eagle,W.O., Gupta,J., Ho,S.-L., Idol,J., Lee-Ilh,S.-O., Legaspi,R., Lim,M., Maduro,Q.L., Maduro,V.B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantiripop,S., Summers,T.J., Thomas,J.W., Thomas,P.J., Tlionson,E.E., Touchman,J.W., Tyrant,J.T., Vogt,J.L., Walker,M.A., Welchthy,K.D. and Green,E.D.				
TITLE	NISC Mouse Sequencing Initiative				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 213308)				
AUTHORS	Green,E.D.				
TITLE	Direct Submission				
JOURNAL	Submitted (22-AUG-2000) NIH Intramural Sequencing Center, 8717				
AUTHORS	Givemont Circle, Galthersburg, MD 20877, USA				
TITLE	Genome Center				
JOURNAL	Center: NIH Intramural Sequencing Center				
AUTHORS	Center code: NISC				
TITLE	Web site: <a href="http://www.nisc.nih.gov">http://www.nisc.nih.gov</a>				
JOURNAL	Contact: <a href="mailto:nisc.mouse@nih.gov">nisc.mouse@nih.gov</a>				
AUTHORS	Project Information				
TITLE	Center project name: sj				
JOURNAL	Center clone name: 157F24				
AUTHORS	Summary Statistics				
TITLE	Sequencing vector: plasmid; n/a; 100% of reads				
JOURNAL	Chemistry: Dye-terminator Big Dye, 100% of reads				
AUTHORS	Assembly program: Phrap; version 0.990319				
TITLE	Consensus quality: 205730 bases at least Q40				
JOURNAL	Consensus quality: 208082 bases at least Q30				
AUTHORS	Consensus quality: 208945 bases at least Q20				
TITLE	Insert size: 202000; agarose-gel				
JOURNAL	Insert size: 216000; pulse-field-gel				
AUTHORS	Insert size: 211508; sum-of-contrigs				
TITLE	Quality coverage: 5.35x in Q20 bases; agarose-gel				
JOURNAL	Quality coverage: 5.00x in Q20 bases; pulse-field-gel				
AUTHORS	Quality coverage: 5.11x in Q20 bases; sum-of-contrigs				
TITLE	NOTE: This is a 'working draft' sequence. It currently				
JOURNAL	consists of 19 contrigs. The true order of the pieces				
AUTHORS	is not known and their order in this sequence record is				
TITLE	arbitrary. Gaps between the contrigs are represented as				
JOURNAL	runs of N, but the exact sizes of the gaps are unknown.				
AUTHORS	This record will be updated with the finished sequence				
TITLE	as soon as it is available and the accession number will				
JOURNAL	be preserved.				
AUTHORS	1				
TITLE	2366: contig of 2366 bp in length				
JOURNAL	2367				
AUTHORS	2466: gap of unknown length				
TITLE	2467				
JOURNAL	6729: contig of 4263 bp in length				
AUTHORS	6730				
TITLE	6829: gap of unknown length				
JOURNAL	6830				
AUTHORS	12022: contig of 5193 bp in length				
TITLE	12023				
JOURNAL	12122: gap of unknown length				
AUTHORS	12123				
TITLE	19872: contig of 7750 bp in length				
JOURNAL	19873				
AUTHORS	19972: gap of unknown length				
TITLE	19973				
JOURNAL	24545: contig of 4573 bp in length				
AUTHORS	24546				
TITLE	24645: gap of unknown length				
JOURNAL	31721: contig of 7076 bp in length				
AUTHORS	31722				
TITLE	31821: gap of unknown length				
JOURNAL	31822				
AUTHORS	37483: contig of 5662 bp in length				
TITLE	37484				
JOURNAL	37583: gap of unknown length				

	*	37584	44430.. contig of 6847 bp in length
	*	44431	44530.. gap of unknown length
	*	44531	53266.. contig of 8736 bp in length
	*	53267	53366.. gap of unknown length
	*	53367	62347.. contig of 8981 bp in length
	*	62348	62447.. gap of unknown length
	*	62448	70761.. contig of 8314 bp in length
	*	70762	70861.. gap of unknown length
	*	70862	81880.. contig of 11019 bp in length
	*	81881	81980.. gap of unknown length
	*	81981	95986.. contig of 14006 bp in length
	*	95987	96086.. gap of unknown length
	*	96087	111352.. contig of 15266 bp in length
	*	111353	111452.. gap of unknown length
	*	111453	121065.. contig of 9613 bp in length
	*	121066	121165.. gap of unknown length
	*	121166	135880.. contig of 14715 bp in length
	*	135881	135980.. gap of unknown length
	*	135981	156578.. contig of 20598 bp in length
	*	156579	156678.. gap of unknown length
	*	156679	182336.. contig of 25558 bp in length
	*	182237	182336.. gap of unknown length
	*	182338	213308.. contig of 30972 bp in length
FEATURES SOURCE	Location/Qualifiers		
	1..213308		
misc_feature	/organism="Mus musculus"		
	/strain="C57BL6/J"		
	/db_xref="taxon:10090"		
	/chromosome="12"		
	/clone="RP23-157F24"		
	/clone_idb="RPCI mouse BAC library 23"		
	1..2366		
	/note="assembly_fragment"		
	2467..6729		
	/note="assembly_fragment"		
	6830..12022		
	/note="assembly_fragment"		
	1213..19872		
	/note="assembly_fragment"		
	19973..24545		
misc_feature	/note="assembly_fragment"		
	24646..31721		
	/note="assembly_fragment"		
	clone_end:sp6		
	vector_side:right"		
	31822..37483		
	/note="assembly_fragment"		
	37584..44430		
	/note="assembly_fragment"		
	44531..53366		
	/note="assembly_fragment"		
	53367..62347		
	/note="assembly_fragment"		
	62448..70761		
	/note="assembly_fragment"		
misc_feature	70862..81880		
	/note="assembly_fragment"		
	81981..95986		
	/note="assembly_fragment"		
	96087..111352		
	/note="assembly_fragment"		
	111453..121065		
	/note="assembly_fragment"		
	121166..135880		
	/note="assembly_fragment"		
	clone_end:r7		
	vector_side:right"		
	135981..156578		
	/note="assembly_fragment"		
	156579..182236		
misc_feature	/note="assembly_fragment"		
	182337..213308		
misc_feature	/note="assembly_fragment"		
	/note="assembly_fragment"		

BASE COUNT 65610 a 41484 c 40972 g 63425 t 1817 others  
ORIGIN  
Query Match 92.0%; Score 18.4; DB 1; Length 213308;  
Best Local Similarity 95.0%; Pred. No. 18;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 TTAGTCTTACTTATTACTT 20  
|||||  
Db 77098 TTAGTCTTACTTATTACTT 77117

RESULT 11  
AC122334/c  
LOCUS 217953 bp DNA linear HTG 20-JUN-2002  
DEFINITION Mus musculus chromosome UNK clone RP23-349H2, WORKING DRAFT  
ACCESSION AC122334  
VERSION AC122334.2 GI:21490480  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 217953)  
AUTHORS McPherson,J.D. and Waterston,R.H.  
TITLE The sequence of Mus musculus clone  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 217953)  
AUTHORS McPherson,J.D. and Waterston,R.H.  
TITLE Direct Submission  
JOURNAL Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
3 (bases 1 to 217953)  
McPherson,J.D. and Waterston,R.H.  
REFERENCE Direct Submission  
JOURNAL Submitted (20-JUN-2002) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
COMMENT On Jun 20, 2002 this sequence version replaced gi:21105195.

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc/index.shtml>  
Contact: [submissions@watson.wustl.edu](mailto:submissions@watson.wustl.edu)  
----- Project Information -----  
Center project name: M\_BA0349H02

----- Summary Statistics -----  
Sequencing vector: M13; 0%  
Sequencing vector: plasmid; 100%  
Chemistry: Dye-primer ET; 0% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 215136 bases at least Q40  
Consensus quality: 215765 bases at least Q30  
Consensus quality: 216021 bases at least Q20  
Insert size: 213000; agarose-fp  
Insert size: 217197; sum-of-ctnigs  
Quality coverage: 17.04 in Q20 bases; agarose-fp  
Quality coverage: 12.14 in Q20 bases; sum-of-ctnigs

----- NOTE: This is a 'working draft' sequence. It currently  
\* consists of 8 ctnigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the ctnigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 8830: contig of 8830 bp in length  
\* 8831 8930: gap of unknown length  
\* 8931 18734: contig of 9804 bp in length

18735 18834: gap of unknown length  
\* 18835 37934: contig of 19100 bp in length  
\* 37935 38034: gap of unknown length  
\* 38035 62372: contig of 24338 bp in length  
\* 62373 62472: gap of unknown length  
\* 62473 62528: contig of 56 bp in length  
\* 62529 62628: gap of unknown length  
\* 62629 92509: contig of 29881 bp in length  
\* 92510 92609: gap of unknown length  
\* 92610 126249: contig of 33640 bp in length  
\* 126250 126349: gap of unknown length  
\* 126350 217953: contig of 91604 bp in length.  
Location/Qualifiers  
1..217953  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosome="UNK"  
/clone="RP23-349H2"  
1..8830  
/note="assembly\_name:Contig34"  
8931..18734  
/note="assembly\_name:Contig35"  
18835..37934  
/note="assembly\_name:Contig36"  
38035..62372  
/note="assembly\_name:Contig37"  
62473..62528  
/note="assembly\_name:Contig17"  
62629..92509  
/note="assembly\_name:Contig38"  
92610..126249  
/note="assembly\_name:Contig39"  
126350..217953  
/note="assembly\_name:Contig40"

BASE COUNT 65772 a 42397 c 42080 g 66952 t 712 others  
ORIGIN  
Query Match 92.0%; Score 18.4; DB 1; Length 217953;  
Best Local Similarity 95.0%; Pred. No. 18;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 TTAGTCTTACTTATTACTT 20  
|||||  
Db 143964 TTAGTCTTACTTATTACTT 143945

RESULT 12  
ALB44536  
LOCUS 241585 bp DNA linear HTG 09-AUG-2002  
DEFINITION Mus musculus chromosome 2 clone RP23-22A15, \*\*\* SEQUENCING IN  
PROGRESS \*\*\*; 11 unordered pieces.  
ACCESSION ALB44536  
VERSION ALB44536.3 GI:22204804  
KEYWORDS HTG; HTGS\_PHASE1.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 241585)  
AUTHORS Plumb,B.  
TITLE Direct Submission  
JOURNAL Submitted (02-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
[humquerry@sanger.ac.uk](mailto:humquerry@sanger.ac.uk) Clone requests: [clonerequest@sanger.ac.uk](mailto:clonerequest@sanger.ac.uk)  
On Aug 11, 2002 this sequence version replaced gi:21555751.

----- Genome Center -----  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [humquerry@sanger.ac.uk](mailto:humquerry@sanger.ac.uk)  
----- Project Information -----  
Center project name: BM22A15  
----- Summary Statistics -----





REFERENCE 1 (bases 1 to 102832)  
AUTHORS Skuce,C.  
TITLE Direct Submission  
JOURNAL Submitted (26-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfacesanger.ac.uk  
Clone requests: clonerequest@sanger.ac.uk  
On Jul 29, 2002 this sequence version replaced gi:22002703.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: zfacesanger.ac.uk  
----- Project Information  
Center project name: dz78P07  
----- Summary Statistics  
Sequencing program: XGAP4; version 4.5  
Sequencing vector: plasmid; 108752; 100% of reads  
Chemistry: Dye-terminator; 12% of reads  
Chemistry: Dye-terminator Big Dye; 87% of reads  
Consensus quality: 102547 bases at least Q40  
Consensus quality: 102595 bases at least Q30  
Consensus quality: 102618 bases at least Q20  
Insert size: 102632; sum-of-contigs  
Insert size: 112120; 3.5% error; agarose-fp  
Quality coverage: 8.47x in Q20 bases; sum-of-contigs Quality  
coverage: 7.98x in Q20 bases; agarose-fp  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 3 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 19949: contig of 19949 bp in length  
\* 19950 20049: gap of 100 bp  
\* 20050 23239: contig of 3190 bp in length  
\* 23240 23339: gap of 100 bp  
\* 23340 102832: contig of 79493 bp in length.  
Location/Qualifiers  
1..102832  
/organism="Danio rerio"  
/db\_xref="taxon:7955"  
/clone="BUSM1-78P07"  
/clone\_lib="BUSM1"  
1..19949  
/note="assembly\_fragment:00162  
fragment\_chain:1  
clone\_end:77  
vector\_side:left"  
20050..23339  
/note="assembly\_fragment:01428  
fragment\_chain:1"  
23340..102832  
/note="assembly\_fragment:01396  
fragment\_chain:1"  
misc\_feature  
misc\_feature  
misc\_feature  
BASE COUNT 32038 a 19328 c 19633 g 31633 t 200 others  
ORIGIN  
Query Match 87.0%; Score 17.4; DB 1; Length 102832;  
Best Local Similarity 94.7%; Pred. No. 72;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 TAGTCTTAGTATTAGTT 20  
|||||  
Db 77546 TAGTCTTAGTATTAGTT 77564  
|||||  
RESULT 15  
LOCUS AL713862 147734 bp DNA linear HTG 10-JUL-2002  
DEFINITION Danio rerio clone BUSM1-68H22, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 11

unordered pieces.  
ACCESSION AL713862  
VERSION AL713862.6 GI:20520522  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_CANCELLED.  
SOURCE zedratish.  
ORGANISM Danio rerio  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Telostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.  
1 (bases 1 to 147734)  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: zfacesanger.ac.uk  
----- Project Information  
Center project name: dz68H22  
----- Summary Statistics  
Sequencing program: XGAP4; version 4.5  
Sequencing vector: plasmid; 108752; 100% of reads  
Chemistry: Dye-terminator; 6% of reads  
Chemistry: Dye-terminator Big Dye; 93% of reads  
Consensus quality: 140645 bases at least Q40  
Consensus quality: 142139 bases at least Q30  
Consensus quality: 143745 bases at least Q20  
Insert size: 146734; sum-of-contigs  
Insert size: 105547; 8.3% error; agarose-fp  
Quality coverage: 15.20x in Q20 bases; sum-of-contigs Quality  
coverage: 26.23x in Q20 bases; agarose-fp  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 11 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 2540: contig of 2540 bp in length  
\* 2541 2640: gap of 100 bp  
\* 2641 85546: contig of 82906 bp in length  
\* 85547 85646: gap of 100 bp  
\* 85647 100008: contig of 14362 bp in length  
\* 100009 100108: gap of 100 bp  
\* 100109 106574: contig of 6466 bp in length  
\* 106575 106674: gap of 100 bp  
\* 106675 126703: contig of 20029 bp in length  
\* 126704 126803: gap of 100 bp  
\* 126804 130027: contig of 3224 bp in length  
\* 130028 130127: gap of 100 bp  
\* 130128 133270: contig of 3143 bp in length  
\* 133271 133370: gap of 100 bp  
\* 133371 133771: contig of 2401 bp in length  
\* 133772 135871: gap of 100 bp  
\* 135872 138584: contig of 2713 bp in length  
\* 138585 138684: gap of 100 bp  
\* 138685 140983: contig of 2299 bp in length  
\* 140984 141083: gap of 100 bp  
\* 141084 147734: contig of 6651 bp in length.  
Location/Qualifiers  
1..147734  
/organism="Danio rerio"  
/db\_xref="taxon:7955"  
/clone="BUSM1-68H22"  
/clone\_lib="BUSM1"  
1..2540  
/note="assembly\_fragment:01940  
misc\_feature  
misc\_feature  
misc\_feature

```

misc_feature      fragment_chain:1"
2641..85546
/note="assembly_fragment:01906
fragment_chain:1"
misc_feature      85647..100008
/note="assembly_fragment:02336
fragment_chain:1"
misc_feature      100109..106574
/note="assembly_fragment:06581
fragment_chain:1"
misc_feature      106675..126703
/note="assembly_fragment:03339.0"
126804..130027
/note="assembly_fragment:03339.1"
130128..133270
/note="assembly_fragment:04734"
133371..135771
/note="assembly_fragment:04895"
135872..138584
/note="assembly_fragment:06490"
138685..140983
/note="assembly_fragment:06532"
141084..147734
/note="assembly_fragment:06663"
BASE COUNT      45653 a 28122 c 28427 g 44527 t 1005 others
ORIGIN

```

```

Query Match      87.0%; Score 17.4; DB 1; Length 147734;
Best Local Similarity 94.7%; Pred. No. 74;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY      2 TACTCTTACTTACTT 20
Db 96725 TACTCTTACTTACTT 96743

```

Search completed: April 6, 2003, 11:41:10  
 Job time : 516.5 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 10:34:07 ; Search time 243 Seconds  
(without alignments)  
1069.066 Million cell updates/sec

Title: US-09-980-265-9

Perfect score: 20

Sequence: 1 ttatgtcttattatgatt 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapept 1.0

Searched: 1994485 seqs, 6494577260 residues

Total number of hits satisfying chosen parameters: 3988970

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl\_NoHTG:\*

1: gb\_ba:\*  
2: gb\_in:\*  
3: gb\_om:\*  
4: gb\_ov:\*  
5: gb\_ph:\*  
6: gb\_pl:\*  
7: gb\_pr:\*  
8: gb\_ro:\*  
9: gb\_sts:\*  
10: gb\_sy:\*  
11: gb\_un:\*  
12: gb\_vl:\*  
13: gb\_vl:\*  
14: em\_ba:\*  
15: em\_fun:\*  
16: em\_hum:\*  
17: em\_in:\*  
18: em\_mu:\*  
19: em\_om:\*  
20: em\_or:\*  
21: em\_ov:\*  
22: em\_pat:\*  
23: em\_ph:\*  
24: em\_pl:\*  
25: em\_ro:\*  
26: em\_sts:\*  
27: em\_un:\*  
28: em\_vl:\*  
29: em\_htg\_hum:\*  
30: em\_htg\_inv:\*  
31: em\_htg\_other:\*  
32: em\_htg\_mus:\*  
33: em\_htg\_pln:\*  
34: em\_htg\_rnd:\*  
35: em\_htg\_mam:\*  
36: em\_htg\_vit:\*  
37: em\_sy:\*  
38: em\_htgo\_hum:\*  
39: em\_htgo\_mus:\*  
40: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	20	100.0	20	5	AX057375 Sequence
2	18.4	92.0	20	5	AX057367 Sequence
3	18.4	92.0	20	5	AX057371 Sequence
4	18.4	92.0	20	5	AX057376 Sequence
5	18.4	92.0	20	5	AX057377 Sequence
6	18.4	92.0	20	5	AX057378 Sequence
7	18.4	92.0	20	5	AX057379 Sequence
8	18.4	92.0	65	5	AX485115 Sequence
9	18.4	92.0	749	7	AF402509 Primula c
10	18.4	92.0	25102	7	AF128394 Arabidops
11	18.4	92.0	35468	2	AF098500 Caenorhab
12	18.4	92.0	45205	2	U41014 Caenorhabd
13	18.4	92.0	105138	8	AF124730 Homo sapi
14	18.4	92.0	153094	8	CNS05TCH
15	18.4	92.0	165920	8	AC010232 Homo sapi
16	18.4	92.0	170892	8	CNS07EES Human chr
17	18.4	92.0	174919	2	AC009370 Drosophill
18	18.4	92.0	175867	2	AC007976 Drosophill
19	18.4	92.0	178628	2	AC009385 Drosophill
20	18.4	92.0	189413	8	AF129075 Homo sapi
21	18.4	92.0	196296	7	ATCHRIV10 Arabidops
22	18.4	92.0	292226	2	AE003520 Drosophill
23	18.4	92.0	324484	8	AE003584 Drosophill
24	18.4	92.0	340000	1	HS21C049 Homo sapi
25	18.4	92.0	349050	1	AP003586 Nostoc sp
26	18	90.0	81332	7	AP004959 Lotus jap
27	17.4	87.0	519	7	AF464739 Coelogyne
28	17.4	87.0	569	7	AF464744 Coelogyne
29	17.4	87.0	570	7	AF503678 Dendrochi
30	17.4	87.0	576	7	AF464745 Coelogyne
31	17.4	87.0	589	7	AF464740 Coelogyne
32	17.4	87.0	600	7	AF464746 Coelogyne
33	17.4	87.0	600	7	AF464750 Coelogyne
34	17.4	87.0	601	7	AF464748 Coelogyne
35	17.4	87.0	602	7	AF464741 Coelogyne
36	17.4	87.0	602	7	AF464749 Coelogyne
37	17.4	87.0	603	7	AF464742 Coelogyne
38	17.4	87.0	603	7	AF464743 Coelogyne
39	17.4	87.0	603	7	AF464751 Coelogyne
40	17.4	87.0	603	7	AF464752 Neogyne g
41	17.4	87.0	603	7	AF464753 Pholidota
42	17.4	87.0	651	1	AB031217 Flavobact
43	17.4	87.0	651	1	AB031221 Flavobact
44	17.4	87.0	1955	8	AB052138 Macaca fa
45	17.4	87.0	7238	5	AX345350 Sequence

## ALIGNMENTS

RESULT 1  
LOCUS AX057375 20 bp DNA linear PAT 17-JAN-2001  
DEFINITION Sequence 9 from Patent W00075304.  
ACCESSION AX057375  
VERSION AX057375.1 GI:12310116  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct.  
artificial sequences.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Bachy,M., Sodoyer,R. and Tranony,E.  
TITLE Immunostimulant oligonucleotide  
JOURNAL Patent: WO 0075304-A 9 14-DEC-2000;  
Aventis Pasteur (FR)  
FEATURES  
Location/Qualifiers

source 1..20  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="Oligonucleotide"  
BASE COUNT 4 a 1 c 3 g 12 t  
ORIGIN

Query Match 100.0%; Score 20; DB 5; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTTATTAGTT 20  
|||||  
Db 1 TTAGTCTTAGTTATTAGTT 20

RESULT 2  
AX057367 AX057367 20 bp DNA linear PAT 17-JAN-2001  
LOCUS  
DEFINITION Sequence 1 from Patent WO0075304.  
ACCESSION AX057367  
VERSION AX057367.1 GI:12310108  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM artificial sequences.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Bachy,M., Sodoyer,R. and Tranoy,E.  
TITLE Immunostimulant oligonucleotide  
JOURNAL Patent: WO 0075304-A 1 14-DEC-2000;  
Aventis Pasteur (FR)  
FEATURES  
source Location/Qualifiers  
1..20  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="Oligonucleotide"

BASE COUNT 5 a 0 c 3 g 12 t  
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;  
Best Local Similarity 95.0%; Pred. No. 6.1e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTTATTAGTT 20  
|||||  
Db 1 TTAGTCTTAGTTATTAGTT 20

RESULT 3  
AX057371 AX057371 20 bp DNA linear PAT 17-JAN-2001  
LOCUS  
DEFINITION Sequence 5 from Patent WO0075304.  
ACCESSION AX057371  
VERSION AX057371.1 GI:12310112  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM artificial sequences.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Bachy,M., Sodoyer,R. and Tranoy,E.  
TITLE Immunostimulant oligonucleotide  
JOURNAL Patent: WO 0075304-A 5 14-DEC-2000;  
Aventis Pasteur (FR)  
FEATURES  
source Location/Qualifiers  
1..20  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="Oligonucleotide"

BASE COUNT 4 a 0 c 3 g 13 t  
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;  
Best Local Similarity 95.0%; Pred. No. 6.1e+02;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTTATTAGTT 20  
|||||  
Db 1 TTAGTCTTAGTTATTAGTT 20

RESULT 4  
AX057376 AX057376 20 bp DNA linear PAT 17-JAN-2001  
LOCUS  
DEFINITION Sequence 10 from Patent WO0075304.  
ACCESSION AX057376  
VERSION AX057376.1 GI:12310117  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM artificial sequences.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Bachy,M., Sodoyer,R. and Tranoy,E.  
TITLE Immunostimulant oligonucleotide  
JOURNAL Patent: WO 0075304-A 10 14-DEC-2000;  
Aventis Pasteur (FR)  
FEATURES  
source Location/Qualifiers  
1..20  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="Oligonucleotide"

BASE COUNT 3 a 1 c 3 g 13 t  
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;  
Best Local Similarity 95.0%; Pred. No. 6.1e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTTATTAGTT 20  
|||||  
Db 1 TTAGTCTTAGTTATTAGTT 20

RESULT 5  
AX057377 AX057377 20 bp DNA linear PAT 17-JAN-2001  
LOCUS  
DEFINITION Sequence 11 from Patent WO0075304.  
ACCESSION AX057377  
VERSION AX057377.1 GI:12310118  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM artificial sequences.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Bachy,M., Sodoyer,R. and Tranoy,E.  
TITLE Immunostimulant oligonucleotide  
JOURNAL Patent: WO 0075304-A 11 14-DEC-2000;  
Aventis Pasteur (FR)  
FEATURES  
source Location/Qualifiers  
1..20  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="Oligonucleotide"

BASE COUNT 3 a 2 c 3 g 12 t  
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;  
Best Local Similarity 95.0%; Pred. No. 6.1e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTTATTAGTT 20  
|||||  
Db 1 TTAGTCTTAGTTATTAGTT 20

RESULT 6  
AX057378

LOCUS AX057378 20 bp DNA linear PAT 17-JAN-2001  
DEFINITION Sequence 12 from Patent WO0075304.  
ACCESSION AX057378  
VERSION AX057378.1 GI:12310119  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Bachy,M., Sodoyer,R. and Tranoy,E.  
TITLE Immunostimulant oligonucleotide  
JOURNAL Patent: WO 0075304-A 12 14-DEC-2000;  
Aventis Pasteur (FR)  
FEATURES  
source location/Qualifiers  
1..20  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="oligonucleotide"  
BASE COUNT 4 a 0 c 4 g 12 t  
ORIGIN  
Query Match 92.0%; Score 18.4; DB 5; Length 20;  
Best Local Similarity 95.0%; Pred. No. 6.1e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TTAGTCTTAGTATTAGTT 20  
||||| ||||||| |||||  
Db 1 TTAGTCTTAGTATTAGTT 20  
||||| ||||||| |||||  
RESULT 7  
LOCUS AX057379 20 bp DNA linear PAT 17-JAN-2001  
DEFINITION Sequence 13 from Patent WO0075304.  
ACCESSION AX057379  
VERSION AX057379.1 GI:12310120  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Bachy,M., Sodoyer,R. and Tranoy,E.  
TITLE Immunostimulant oligonucleotide  
JOURNAL Patent: WO 0075304-A 13 14-DEC-2000;  
Aventis Pasteur (FR)  
FEATURES  
source location/Qualifiers  
1..20  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="oligonucleotide"  
BASE COUNT 3 a 1 c 4 g 12 t  
ORIGIN  
Query Match 92.0%; Score 18.4; DB 5; Length 20;  
Best Local Similarity 95.0%; Pred. No. 6.1e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TTAGTCTTAGTATTAGTT 20  
||||| ||||||| |||||  
Db 1 TTAGTCTTAGTATTAGTT 20  
||||| ||||||| |||||  
RESULT 8  
LOCUS AX485115 65 bp DNA linear PAT 16-AUG-2002  
DEFINITION Sequence 2415 from Patent WO02053728.  
ACCESSION AX485115  
VERSION AX485115.1 GI:22319399  
KEYWORDS  
SOURCE Candida albicans.  
ORGANISM Candida albicans.  
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; mitosporic Saccharomycetales; Candida.

REFERENCE 1  
AUTHORS Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.  
TITLE Gene disruption methodologies for drug target discovery  
JOURNAL Patent: WO 02053728-A 2415 11-JUN-2002;  
Elitra Pharmaceuticals, Inc. (US)  
FEATURES  
source location/Qualifiers  
1..65  
/organism="Candida albicans"  
/db\_xref="taxon:5476"  
BASE COUNT 37 a 6 c 16 g 6 t  
ORIGIN  
Query Match 92.0%; Score 18.4; DB 5; Length 65;  
Best Local Similarity 95.0%; Pred. No. 4.5e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TTAGTCTTAGTATTAGTT 20  
||||| ||||||| |||||  
Db 46 TTAGTCTTAGTATTAGTT 27  
||||| ||||||| |||||  
RESULT 9  
LOCUS AF402509 749 bp DNA linear PLN 05-NOV-2001  
DEFINITION Primula cicutarillifolia ribosomal protein L16 (rp116) gene, intron  
sequence; chloroplast gene for chloroplast product.  
ACCESSION AF402509  
VERSION AF402509.1 GI:16660214  
KEYWORDS  
SOURCE Primula cicutarillifolia.  
ORGANISM Chloroplast Primula cicutarillifolia  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asterales; Ericales; Primulaceae; Primula.  
1 (bases 1 to 749)  
Mast,A.R., Kelso,S., Richards,A.J., Lang,D.J., Feller,D.M.S. and  
Conti,E.  
Phylogenetic Relationships in Primula L. and Related Genera  
(Primulaceae) based on Noncoding Chloroplast DNA  
Int. J. Plant Sci. 162 (6), 1381-1400 (2001)  
2 (bases 1 to 749)  
Mast,A.R., Kelso,S., Richards,A.J., Lang,D.J., Feller,D.M.S. and  
Conti,E.  
Direct Submission  
TITLE Submitted (25-JUL-2001) Institute for Systematic Botany, University  
of Zurich, Zollikerstrasse 107, Zurich 8008, Switzerland  
JOURNAL  
AUTHORS  
FEATURES  
source location/Qualifiers  
1..749  
/organism="Primula cicutarillifolia"  
/organella="plastid:chloroplast"  
/db\_xref="taxon:170920"  
<1..>749  
/gene="rp116"  
/note="ribosomal protein L16"  
<1..>749  
/gene="rp116"  
BASE COUNT 305 a 110 c 129 g 205 t  
ORIGIN  
Query Match 92.0%; Score 18.4; DB 7; Length 749;  
Best Local Similarity 95.0%; Pred. No. 2.4e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TTAGTCTTAGTATTAGTT 20  
||||| ||||||| |||||  
Db 47 TTGGTCTTAGTATTAGTT 28  
||||| ||||||| |||||  
RESULT 10  
LOCUS T25H8 25102 bp DNA linear PLN 03-MAR-1999  
DEFINITION Arabidopsis thaliana BAC T25H8.  
ACCESSION AF128394



Db 23783 TTAGTCTTACTGACTTACTT 23802

|||||

RESULT 11 35468 bp DNA linear INV 22-MAY-2002  
AF098500  
LOCUS  
DEFINITION Caenorhabditis elegans cosmid F43B10, complete sequence.  
ACCESSION AF098500  
VERSION AF098500.1 GI:3786411  
KEYWORDS HTG.  
SOURCE  
ORGANISM Caenorhabditis elegans.  
Eukaryota; Metazoa; Nematoda; Chromodorea; Rhabditida;  
Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.

REFERENCE  
AUTHORS Waterston, R.  
TITLE 1 (bases 1 to 35468)

REFERENCE  
AUTHORS Waterston, R.  
TITLE Genome sequence of the nematode C. elegans: a platform for  
investigating biology. The C. elegans Sequencing Consortium  
Science 282 (5396), 2012-2018 (1998)

REFERENCE  
MEDLINE 99069613  
PUBMED 9851916

REFERENCE  
AUTHORS Fulton, R., Hawkins, J. and Rohlfing, T.  
TITLE 2 (bases 1 to 35468)

JOURNAL The sequence of C. elegans cosmid F43B10

REFERENCE  
AUTHORS Unpublished (2001)

JOURNAL 3 (bases 1 to 35468)

REFERENCE  
AUTHORS Waterston, R.  
TITLE Direct Submission

JOURNAL Submitted (13-OCT-1998) Department of Genetics, Washington  
University, 444 Forest Park Avenue, St. Louis, Missouri 63108, USA

REFERENCE  
AUTHORS Waterston, R.  
TITLE 4 (bases 1 to 35468)

JOURNAL Direct Submission  
Submitted (03-JUL-2001) Department of Genetics, Washington  
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.  
Louis, MO 63110, USA

REFERENCE  
AUTHORS 5 (bases 1 to 35468)

JOURNAL Waterston, R.  
TITLE Direct Submission

JOURNAL Submitted (22-MAY-2002) Department of Genetics, Washington  
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.  
Louis, MO 63110, USA

COMMENT  
Submitted by:

Genome Sequencing Center  
Department of Genetics, Washington University  
St. Louis, MO 63110, USA, and  
Sanger Centre, Hinxton Hall  
Cambridge CB10 1RQ, England  
email: tw@nematoe.wustl.edu and jee@sanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone.  
It may be shorter because we only sequence overlapping sections  
once, or longer because we provide a small overlap between  
neighboring submissions.

This sequence was finished as follows unless otherwise noted: all  
regions were double stranded, sequenced with an alternate chemistry  
or covered by high quality data (i.e., phred quality >= 30); an  
attempt was made to resolve all sequencing problems, such as  
compressions and repeats; all regions were covered by sequence from  
more than one m13 subclone.

For a graphical representation of this cosmid sequence and its  
analysis see:

http://www.wormbase.org/db/seq/sequence?name=F43B10;class=Sequence

NEIGHBORING COSMID INFORMATION

The 5' cosmid is C06G1, 14200 bp overlap; the 3' cosmid is F22H10,  
200 bp overlap. Actual start of this cosmid is at base position 197  
of F43B10, actual end is at 3375 of F22H10.

NOTES:

Coding sequences below are the result of integration and manual  
review of the following data: computer analysis using the program  
GeneFinder (P. Green and L. Hillier, personal communication), the  
large scale EST projects of Yuji Kohara  
(http://www.ddb.jni.ac.jp/c-elegans/html/CE\_INDEX.html) and The C.  
elegans ORFome cloning project (http://wormfdb.fdel.harvard.edu/),  
similarity to other proteins from Blastx analyses  
(http://blast.wustl.edu/), sequence conservation with C. briggsae  
using Jim Kent's WABA alignment program (Genome Research  
10:1115-1125, 2000), individual C. elegans Genbank submissions,  
and personal communications with C. elegans researchers. tRNAs  
are predicted using the program tRNAscan-SE ( Lowe, T. M. and  
Eddy, S.R., 1997, Nucl. Acids. Res., 25, 955-964).

FEATURES  
SOURCE

1..35468

/organism="Caenorhabditis elegans"

/strain="Bristol N2"

/db\_xref="taxon:6239"

/chromosome="X"

/clone="F43B10"

/complement(20574..22725)

/gene="F43B10.1"

/note="for a graphical representation of this gene see:  
http://www.wormbase.org/db/seq/sequence?name=F43B10.1;clas  
s=Sequence"

complement(join(20574..20880,21108..21251,21323..21503,  
21858..21942,22322..22367,22463..22725))

/gene="F43B10.1"

/note="coded for by the following C. elegans cDNAs:  
yk590f2.5"

/codon\_start=1

/product="Hypothetical protein F43B10.1"

/protein\_id="AAC67400.2"

/db\_xref="GI:14369812"

/translation="MAILHLAPRLGDIIFLCVSIFSLVFLSLYLHSHHARRLI  
IFISVFRMLSLIMPPEPPPNKVVVGRPLTAYELNNAIFPMHITMARGRKRL  
KVRESIRFCHDKIVKMLSPDRSSSGAEVLSIEDDDKMPKSYTKPMEKLD  
LSHGPSDLRSDLPALITPPRKNNVNDLRSPRIETTSADRNVLERLRRKSIES  
PRTFLNGLAGIESLPAYIAKKDNVNSQLTSMWITFTSPHRRERMGGGGDD  
AIRKMSQEVKKRRRIEESGGGGRGVNAWIMAKRT"

/complement(24967..34263)

/gene="F43B10.2"

/note="for a graphical representation of this gene see:  
http://www.wormbase.org/db/seq/sequence?name=F43B10.2;clas  
s=Sequence"

complement(join(24967..25126,25277..25463,25509..25533,  
25750..25963,26181..26287,26340..26520,26565..26740,  
26795..26853,26901..27089,27340..27450,27497..27727,  
27887..27950,29903..30015,30752..31138,31218..31327,  
32093..32175,32472..32631,32674..32806,33844..33953,  
33996..34135,34182..34263))

/gene="F43B10.2"

/note="contains similarity to human ALR (GB:AF003386);  
coded for by the following C. elegans cDNAs: yk12e10.3,  
yk12e10.5, yk350b9.3, yk350b9.5"

/codon\_start=1

/product="Hypothetical protein F43B10.2"

/protein\_id="AAC67399.1"

/db\_xref="GI:3786412"

/translation="MISPSKFEVGEALROKDVIVFSAKEDNERENVYLLGPTS  
SOKSELIDFLCNFYGVSDMDQRPHISNEKFNSETPKRPQCYFNFETKMSVRVIL  
DMACGDYDGEFSLILHKMLNNKMKRLVIAIVFSLRMSHHEDELOOVSNTL  
PEHVRNSIVVITFASDGSRTPEPLRRGLSECPFTINTSCCTKQWEDLNDHR  
RYKMSVNOEDSEMEKQVETTPITISGLPEPDDGICVQESGONSRSRSVSSKSTV  
IEVQSSKQOVQGRTESEPLPEPLPLSNPAPKPPPPPPPIIRPPPIIPASPPR  
PPASEPVGIPYSEKQHOQOLRPTTASPPSPAPPAALITDTSTTNALIN  
AMFQLOKOSVRYDVPYKSDRSRFTQTCQSLPPAGVGRHSVPDVHAYDETT  
PPVYLHVADRYAVYVWNEHPTNSTQVSHSPRSQVGRDARLSRDRLRVYDAO  
SSPADSOSEELRRMYSGRSGSVKPTGVTAARYVDTNRNRYTAGEOARSTSTFT  
NDRKMRSSSEPSSEYITINDPYISGRVYVSGYVETFRRSRTSLHSGVPELILHS  
NKYDQQLTDQONEMERKSRIDYERRSRROSGSRLYIDRQISPERKDDYSGTY

CDS

gene





trna complement(8832. .8903)  
/product="trna-asp"  
/complement(9869. .9940)  
/product="trna-asp"  
12433. .13037  
/gene="C06G1.2"  
/note="for a graphical representation of this gene see:  
http://www.wormbase.org/db/seq/sequence?name=C06G1.2;class  
=Sequence"  
join(12433. .12799,12856. .13037)  
/gene="C06G1.2"  
/codon-start=1  
/product="Hypothetical protein C06G1.2"  
/protein\_id="AAK32945.1"  
/db\_xref="GI:13606083"  
/translation="MNFALNSALFFVLLASACFOMYDDVIGAVVSGSGVEIKASP  
TEATTTTVEETTTTSTTTEETTTTETTTTETTTTETTTTETTTTETTTTETTT  
TTTAAATTTTTEETTTTETTTTETTTTETTTTETTTTETTTTETTTTETTTTET  
TPEQHHPGQKHVQGFQGGGF"  
23904. .27748  
/gene="C06G1.1"  
/note="for a graphical representation of this gene see:  
http://www.wormbase.org/db/seq/sequence?name=C06G1.1;class  
=Sequence"  
join(23904. .23967,24026. .24183,24567. .24665,24716. .24807,  
24973. .25153,25224. .25338,25698. .25769,25840. .26091,  
26292. .26419,26517. .26642,26692. .26859,27644. .27748)  
/gene="C06G1.1"  
/note="coded for by the following C. elegans cDNAs:  
yk361c9.3, yk361c9.5, yk623g9.3, yk623g9.5, yk110e5.3,  
yk110e5.5"  
/codon-start=1  
/product="Hypothetical protein C06G1.1"  
/protein\_id="AAK32948.1"  
/db\_xref="GI:13606086"  
/translation="MIVYLVITLISVNTQQOISLNSLFAEGTSPGVTRINKALS  
IVSQLDVAVKEMNSDGLFENISAPLTNOYRFLRSKIIISFELYFDSMSIVPQK  
GVSMTGINIVNTLISLFTLTOGEITGNVLTQVLTGKIVLWLVGDLNLTDLI  
TCKVAANNLLOFQSDASALLANYIPHOMNIRQTEVQVSPFAELVPLNSRIM  
TPLSALFEQEFINVALGPEVSSOATLTHRNSFECILROGRRLNDFLPLRSP  
LOVPTADSGHIDFVLSYNTMSLTFMDDXKRDYKDSRATNSALVGLKDCGP  
GDICAGTLFPLGTRFPGGEVYIKSHVYSTRVYLRKNNMTYIDSRDAVQOQDRS  
RRFLTASNAEVLKDKPSFKDFVLHGLERIDKFVSDVASLVDISDESLFVNALT  
ELIINEMAKRLKSGIQLPIFDYTOQAANVSIIDDMHLISADFCATDKCATADNKO  
TDVYDYTVQV"  
complement(25366. .25437)  
/product="trna-his"  
/complement(29258. .45019)  
/gene="C06G1.5"  
/note="for a graphical representation of this gene see:  
http://www.wormbase.org/db/seq/sequence?name=C06G1.5;class  
=Sequence"  
complement(join(29258. .29318,29679. .29738,29789. .30127,  
30178. .30319,30371. .30471,30941. .31228,31277. .31446,  
31494. .31590,31952. .32254,32779. .32831,42656. .45019))  
/gene="C06G1.5"  
/note="coded for by the following C. elegans cDNAs:  
yk886h07.3, yk299.5, yk49b8.5, yk161h12.5, yk253a7.5,  
yk886h07.5, yk49b8.3, yk161h12.3, yk253a7.3"  
/codon-start=1  
/product="Hypothetical protein C06G1.5"  
/protein\_id="AAK32946.2"  
/db\_xref="GI:18652581"  
/translation="MYHLDDYFARFNOPSSSSVOFPESEIRYGGSLATTTTGAEN  
DPQTTTTRSSOSADNRKSAFHQPTTISDHIPEPMNQPLETAENPFSQORRASTT  
PWSHSEYHSHSLNLSVPSFSEHLREKSPAGSLALDLISYPAQHSVQSPVQAY  
IDRIHTDRAPTRTSTYQSSSENPLEYISIQRYVNPVRSSTYLVQVNVNGGAPY  
QOYQUTITSMGPTSTHREYHNSTPEYOHSSSLGCPPLTPRESSLOOQOAAHP  
FENRSEFHHPPEPPSPSSSSSHOFPONLESDDLSPASFTTGSMTLPRRQAOLPISK  
RLIDDNKSRSTSEMSKSFEOQNRFONLESDDLSPASFTTGSMTLPRRQAOLPISK  
SSSNTYEMNSSLIGRAVERRAEAPDQKSGALDNFWSPTLSSRPSSTLQIPLSLTA  
AERLOMLQEPIDTEKRFQRYVSGQGAARAEMLREPSVETKPRPLNPAPPSNS  
FLIPSFELDNAVNELSRTVGNFTGGTTFGVSVPAPVYHNSINRRVGGCGRPTTGE

HAFSPPLENSKSPALSLHRKLSPSPTEVSESSSYQIPTGLPHPKPHKHNKEOLH  
Query Match 92.0%; Score 18.4; DB 2; Length 45205;  
Best Local Similarity 95.0%; Pred. No. 80;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTAGTTCCTAGTTATTAGTT 20  
|||||  
Db 38063 TTAGTTCCTAGTTATTAGTT 38082

RESULT 13  
AF124730 105138 bp DNA linear PRI 13-APR-2001  
LOCUS AF124730  
DEFINITION Homo sapiens chromosome 21 clone PAC N2184 map 21q22.1, complete  
sequence.  
ACCESSION AF124730  
VERSION AF124730.2 GI:13621228  
KEYWORDS HTG.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS  
1 (bases 1 to 105138)  
Hattori,M., Fujiyama,A., Taylor,T.D., Matanabe,H., Yada,T.,  
Park,H.-S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.-K., Soeda,E.,  
Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,  
Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,  
Patterson,D., Reichwald,K., Rump,A., Schillhabel,M.B., Schudy,A.,  
Zimmermann,W., Rosenthal,A., Kudoh,J., Kawasaki,K., Asakawa,S.,  
Shintani,A., Sasaki,T., Nagamine,K., Mitsuyma,S.,  
Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordstiek,G.,  
Hornischer,K., Brandt,P., Scherfe,M., Schoen,O., Desario,A.,  
Reichelt,J., Kauer,G., Bloeker,H., Ramser,J., Beck,A., Klages,S.,  
Hennig,S., Riesselmann,L., Dagand,E., Haaf,T., Wehmeyer,S.,  
Borzym,K., Gardiner,K., Mizel,D., Francis,F., Lehrach,H.,  
Reinhardt,R. and Yaspo,M.Laure.  
The DNA sequence of human chromosome 21  
Nature 405 (6784), 311-319 (2000)  
10830953  
2 (bases 1 to 105138)  
Taudien,S., Dagand,E., Delabar,J., Ortl,R., Nordstiek,G.,  
Drescher,B., Weber,J., Schattevoy,R., Yaspo,M.-L. and Rosenthal,A.  
Direct Submission  
Submitted (29-JAN-1999) Genome Analysis, Institut of Molecular  
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany  
3 (bases 1 to 105138)  
Taudien,S., Dagand,E., Delabar,J., Ortl,R., Nordstiek,G.,  
Drescher,B., Weber,J., Schattevoy,R., Yaspo,M.-L. and Rosenthal,A.  
Direct Submission  
Submitted (13-APR-2001) Genome Analysis, Institute of Molecular  
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany  
On Apr 13, 2001 this sequence version replaced gi:4262557.

COMMENT  
FEATURES  
source  
1. 105138  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22.1"  
/clone="PAC N2184"

BASE COUNT 30268 a 20684 c 20685 g 33501 t  
ORIGIN

Query Match 92.0%; Score 18.4; DB 8; Length 105138;  
Best Local Similarity 95.0%; Pred. No. 64;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTAGTTCCTAGTTATTAGTT 20  
|||||  
Db 20497 TTAGTTCCTAGTTATTAGTT 20478

```

RESULT 14
CNS05TCH      153094 bp      DNA      linear      PRI 08-JUN-2001
LOCUS          Human chromosome 14 DNA sequence BAC R-241EJ3 of library RPCI-11
DEFINITION    from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION     AL355098
VERSION       AL355098.3      GI:13677194
KEYWORDS      HTG; HTGS; ACTIVEFLN.
SOURCE        human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 153094)
AUTHORS       Hellig, R., Petit, J.L., Vico, V., Dasilva, C., Robert, C., Mincker, P.,
              Brothier, P., Catolico, L., Barbe, V., Pelletier, E., Artiguenave, F.,
              Levy, M., Eckenberg, R., Bruls, T., deBardins, V., Cruaud, C.,
              Gyaray, G., Saurin, W. and Weissbach, J.
              Sequencing of the human chromosome 14
              Unpublished
              2 (bases 1 to 153094)
TITLE         Direct Submission
AUTHORS       Genoscope.
JOURNAL       Submitted (07-JUN-2001) Genoscope - Centre National de Sequencage :
              BP 191 91006 Evry cedex - FRANCE (E-mail : segref@genoscope.cns.fr
              Web : www.genoscope.cns.fr)
              On Apr 19, 2001 this sequence version replaced gi:9213458.
COMMENT       ----- Genome Center
              Center: Genoscope / Centre National de Sequencage
              Web site: http://www.genoscope.cns.fr/
              Contact: Segref@genoscope.cns.fr

              The following BAC sequence is oriented from the T7 to the SP6 end.
              Upstream BAC (overlapping the T7 end) : C-2506J14
              Downstream BAC (overlapping the SP6 end) : R-76E12 (AC-AL163872).
              ----- Summary Statistics
              Assembly program: Phrap; version 2.0
              Quality coverage: 8.79x in Q20 bases; sum-of-contigs

              -----
              Overall quality chart :
              Range : bases
              0 - 9 :
              10 - 19 :
              20 - 29 :
              30 - 39 : 34
              40 - 49 : 1110
              50 - 59 : 4764
              60 - 69 : 4337
              70 - 79 : 8503
              80 - 89 : 30075
              90 - 99 : 104271

              -----
              Percentage of bases with a quality value >= 40 : 99 %.
              Location/Qualifiers
                1. 153094
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /chromosome="14"
                  /clone="R-241EJ3"
                  /clone_1lb="RPCI-11"
BASE COUNT    43932 a 31270 c 32456 g 45436 t
ORIGIN
Query Match          92.0%: Score 18.4; DB 8; Length 153094;
Best Local Similarity 95.0%: Pred. No. 58;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 TTAGTCTTAGTTATTTAGTT 20
        ||||| ||||| ||||| |||||
DB 31203 TTAGTCTTAATTTATTTAGTT 31222

```

```

RESULT 15
AC010232/c     165920 bp      DNA      linear      PRI 26-FEB-2002
LOCUS          Homo sapiens chromosome 5 clone CTC-316C21, complete sequence.
DEFINITION    AC010232
ACCESSION     AC010232
VERSION       AC010232.10      GI:18921277
KEYWORDS      HTG.
SOURCE        Homo sapiens.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 165920)
AUTHORS       DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL       Direct Submission
TITLE         Unpublished
AUTHORS       2 (bases 1 to 165920)
JOURNAL       DOE Joint Genome Institute.
TITLE         Direct Submission
AUTHORS       Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint
              Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
              3 (bases 1 to 165920)
              DOE Joint Genome Institute and Stanford Human Genome Center.
              Direct Submission
              Submitted (26-FEB-2002) DOE Joint Genome Institute, 2800 Mitchell
              Drive, Walnut Creek, CA 94598, USA
              On Feb 26, 2002 this sequence version replaced gi:15290321.
              Draft Sequence Produced by DOE Joint Genome Institute
              www.jgi.doe.gov
              Finishing Completed at Stanford Human Genome Center
              www.shgc.stanford.edu
              Quality: Phrap Quality >=40 99.9% of Sequence;
              Estimated Total Number of Errors is 0.1.
              Location/Qualifiers
                1. 165920
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /chromosome="5"
                  /clone="CTC-316C21"
BASE COUNT    45866 a 32687 c 34867 g 52500 t
ORIGIN
Query Match          92.0%: Score 18.4; DB 8; Length 165920;
Best Local Similarity 95.0%: Pred. No. 57;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 TTAGTCTTAGTTATTTAGTT 20
        ||||| ||||| ||||| |||||
DB 50747 TTAGTTATTTAGTTATTTAGTT 50728

```

Search completed: April 6, 2003, 11:53:21  
Job time : 336 secs

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 11:21:32 ; Search time 151 Seconds  
(without alignments)  
298.278 Million cell updates/sec

Title: US-09-980-265-9

Perfect score: 1  
Sequence: 1 ttagtcttattattattagt 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_101002:.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	22	AAF28880
2	19	95.0	20	22	AAF28871
3	18.4	92.0	20	22	AAF28872
4	18.4	92.0	20	22	AAF28876
5	18.4	92.0	20	22	AAF28881
6	18.4	92.0	20	22	AAF28882
7	18.4	92.0	20	22	AAF28883
8	18.4	92.0	20	22	AAF28884
9	18.4	92.0	749	24	ABQ21728

C	10	18.4	92.0	749	24	ABQ21729	Oligonucleotide fo
	11	17.4	87.0	7238	24	ABL32448	Human immune syste
	12	17.4	87.0	9507	24	ABN80243	Human chemically m
	13	16.8	84.0		20	AAF28873	Immunostimulatory
	14	16.8	84.0		20	AAF28874	Immunostimulatory
	15	16.8	84.0		20	AAF28875	Immunostimulatory
	16	16.8	84.0		20	AAF28876	Immunostimulatory
	17	16.8	84.0		20	AAF28877	Immunostimulatory
	18	16.8	84.0		20	AAF28878	Immunostimulatory
	19	16.8	84.0		20	AAF28879	Immunostimulatory
	20	16.8	84.0		20	AAF28880	Immunostimulatory
	21	16.8	84.0		20	AAF28881	Immunostimulatory
	22	16.8	84.0		20	AAF28882	Immunostimulatory
	23	16.8	84.0		20	AAF28883	Immunostimulatory
	24	16.8	84.0		20	AAF28884	Immunostimulatory
C	25	16.8	84.0	902	24	ABQ22435	Oligonucleotide fo
	26	16.8	84.0	1069	21	AAAC6432	Arabidopsis thalia
	27	16.8	84.0	1071	21	AAAC3993	Arabidopsis thalia
	28	16.8	84.0	1240	21	AAAC45636	Arabidopsis thalia
	29	16.8	84.0	1242	21	AAAC40226	Arabidopsis thalia
C	30	16.8	84.0	1818	22	AAH13793	Human cDNA sequenc
	31	16.8	84.0	4741	23	ABL27984	Human immune syste
	32	16.8	84.0	5453	24	ABL70450	Chemically treated
	33	16.8	84.0	5453	24	AA561413	Human gene regulat
	34	16.8	84.0	5453	24	ABK31481	Signal transductio
	35	16.8	84.0	6759	24	ABL32739	Human immune syste
	36	16.8	84.0	7309	23	ABL29678	Drosophila melanog
	37	16.8	84.0	7353	24	ABL32072	Human immune syste
	38	16.8	84.0	7353	24	AAAD28362	Human chemically t
	39	16.8	84.0	10682	24	ABN80085	Human chemically m
	40	16.8	84.0	11029	22	AA546414	Tumour suppressor
	41	16.8	84.0	11670	24	AA546414	Chemically treated
	42	16.8	84.0	13377	22	AA546476	Tumour suppressor
C	43	16.8	84.0	13377	24	ABL33463	Human immune syste
	44	16.8	84.0	13798	23	ABL8500	Drosophila melanog
	45	16.8	84.0	16258	24	ABL70376	Chemically treated

#### ALIGNMENTS

RESULT 1	
AAF28880	AAF28880 standard; DNA: 20 BP.
XX	
AC	AAF28880;
XX	
DT	09-MAY-2001 (first entry)
XX	
DE	Immunostimulatory oligonucleotide #9 as vaccine adjuvant.
XX	
KW	Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KW	prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW	cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KW	phosphorothioate; ss.
XX	
OS	Synthetic.
XX	
FH	Key
FT	modified_base
FT	1..20
FT	Location/Qualifiers
FT	/*tag= a
FT	/note= "contain phosphorothioate internucleotide linkages"
XX	
FN	WO200075304-A1.
XX	
XX	14-DEC-2000.
PD	
XX	
XX	08-JUN-2000; 2000WO-FR01566.
PF	
XX	
XX	08-JUN-1999; 99FR-0007457.
PR	
PR	06-AUG-1999; 99FR-0010378.
XX	

```
PA (AVET ) AVENTIS PASTEUR.
XX
XX Bachy M, Sodoyer R, Trannoy E;
XX
DR WPI: 2001-041317/05.
XX
PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
PT vaccines for human use, induce lymphocyte proliferation and cytokine
PT secretion -
XX
PS Example 4; Page 13; 30pp; French.
XX
CC Oligonucleotides AAF28872-AAF2887 represent specific examples of an
CC immunostimulatory oligonucleotide of the invention which contains at
CC least one sequence 5'-TTN1N2TT-3' where N1 and N2 are A, T, C or G. The
CC oligonucleotides do not contain any CG dinucleotides in which C are
CC unmethylated. The oligonucleotides are used as human immunostimulants
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
CC They induce proliferation of human lymphocytes, induce secretion of
CC cytokines, especially Interleukin-10 or Interferon-gamma and increase
CC expression of the CD86 activation marker or the CD25 cytokine receptor
CC on human B lymphocytes.
XX
SQ Sequence 20 BP; 4 A; 1 C; 3 G; 12 T; 0 other;

Query Match      100.0%; Score 20; DB 22; Length 20;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTTAACTT 20
   ||||||||||||||||
DB 1 TTAGTCTTACTTAACTT 20

RESULT 2
AAF28871
ID AAF28871 standard; DNA; 20 BP.
XX
AC AAF28871;
XX
DT 09-MAY-2001 (first entry)
XX
DE Generic immunostimulatory oligonucleotide as vaccine adjuvant.
XX
KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor; ss.
XX
OS Synthetic.
XX
PN WO200075304-A1.
XX
PD 14-DEC-2000.
XX
PE 08-JUN-2000; 2000WO-FR01566.
XX
PR 08-JUN-1999; 99FR-0007457.
PR 06-AUG-1999; 99FR-0010378.
XX
PA (AVET ) AVENTIS PASTEUR.
PI Bachy M, Sodoyer R, Trannoy E;
XX
DR WPI: 2001-041317/05.
XX
PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
PT vaccines for human use, induce lymphocyte proliferation and cytokine
PT secretion -
XX
PS Claim 8; Page 17; 30pp; French.
XX
CC This sequence represents a generic example of an immunostimulatory
CC oligonucleotide of the invention which contains at least one sequence
```

```
CC 5'-TTN1N2TT-3' where N1 and N2 are A, T, C or G. The oligonucleotides
CC do not contain any CG dinucleotides in which C are unmethylated.
CC The oligonucleotides are used as human immunostimulants and as adjuvants
CC in therapeutic and prophylactic vaccines for human use. They induce
CC proliferation of human lymphocytes, induce secretion of cytokines,
CC especially Interleukin-10 or Interferon-gamma and increase expression
CC of the CD86 activation marker or the CD25 cytokine receptor on human B
CC lymphocytes.
XX
SQ Sequence 20 BP; 3 A; 1 C; 3 G; 12 T; 1 other;

Query Match      95.0%; Score 19; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 33;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTTAACTT 20
   ||||||||||||||||
DB 1 TTAGTCTTACTTAACTT 20

RESULT 3
AAF28872
ID AAF28872 standard; DNA; 20 BP.
XX
AC AAF28872;
XX
DT 09-MAY-2001 (first entry)
XX
DE Immunostimulatory oligonucleotide #1 as vaccine adjuvant.
XX
KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KW phosphorothioate; ss.
XX
OS Synthetic.
XX
PN WO200075304-A1.
XX
PD 14-DEC-2000.
XX
PE 08-JUN-2000; 2000WO-FR01566.
XX
PR 08-JUN-1999; 99FR-0007457.
PR 06-AUG-1999; 99FR-0010378.
XX
PA (AVET ) AVENTIS PASTEUR.
PI Bachy M, Sodoyer R, Trannoy E;
XX
DR WPI: 2001-041317/05.
XX
PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
PT vaccines for human use, induce lymphocyte proliferation and cytokine
PT secretion -
XX
PS Example 4; Page 13; 30pp; French.
XX
CC Oligonucleotides AAF28872-AAF2887 represent specific examples of an
CC immunostimulatory oligonucleotide of the invention which contains at
CC least one sequence 5'-TTN1N2TT-3' where N1 and N2 are A, T, C or G. The
CC oligonucleotides do not contain any CG dinucleotides in which C are
CC unmethylated. The oligonucleotides are used as human immunostimulants
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
CC They induce proliferation of human lymphocytes, induce secretion of
CC cytokines, especially Interleukin-10 or Interferon-gamma and increase
CC expression of the CD86 activation marker or the CD25 cytokine receptor
```

```
CC on human B lymphocytes.
XX
SQ Sequence 20 BP; 5 A; 0 C; 3 G; 12 T; 0 other;
Query Match 92.0%; Score 18.4; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 59;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTTATTAGTT 20
   ||||| ||||| ||||| |||||
Db 1 TTACTTATTACTTATTAGTT 20

RESULT 4
AAF28876
ID AAF28876 standard; DNA: 20 BP.
XX
AC AAF28876;
XX
DT 09-MAY-2001 (first entry)
XX
DE Immunostimulatory oligonucleotide #5 as vaccine adjuvant.
XX
KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KW phosphorothioate; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1..20
FT /*tag= a
FT /note= "contain phosphorothioate internucleotide
links"
XX
XX WO200075304-A1.
XX
XX 14-DEC-2000.
XX
XX 08-JUN-2000; 2000WO-FR01566.
XX
XX 08-JUN-1999; 99FR-0007457.
XX
XX 06-AUG-1999; 99FR-0010378.
XX
XX (AVET ) AVENTIS PASTEUR.
XX
XX
XX Bachy M, Sodoyer R, Trannoy E;
XX
XX WPI; 2001-041317/05.
XX
XX
XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
XX PT vaccines for human use, induce lymphocyte proliferation and cytokine
XX secretion -
XX
XX Example 4; Page 13; 30pp; French.
XX
XX Oligonucleotides AAF28872-AAF28887 represent specific examples of an
XX immunostimulatory oligonucleotide of the invention which contains at
XX least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The
XX oligonucleotides do not contain any CG dinucleotides in which C are
XX unmethylated. The oligonucleotides are used as human immunostimulants
XX and as adjuvants in therapeutic and prophylactic vaccines for human use.
XX They induce proliferation of human lymphocytes, induce secretion of
XX cytokines, especially interleukin-10 or interferon-gamma and increase
XX expression of the CD86 activation marker or the CD25 cytokine receptor
XX on human B lymphocytes.
XX
XX Sequence 20 BP; 4 A; 0 C; 3 G; 13 T; 0 other;
SQ
Query Match 92.0%; Score 18.4; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 59;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 TTAGTCTTACTTATTAGTT 20
   ||||| ||||| ||||| |||||
Db 1 TTAGTTTACTTATTAGTT 20

RESULT 5
AAF28881
ID AAF28881 standard; DNA: 20 BP.
XX
AC AAF28881;
XX
DT 09-MAY-2001 (first entry)
XX
DE Immunostimulatory oligonucleotide #10 as vaccine adjuvant.
XX
KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KW phosphorothioate; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1..20
FT /*tag= a
FT /note= "contain phosphorothioate internucleotide
links"
XX
XX WO200075304-A1.
XX
XX 14-DEC-2000.
XX
XX 08-JUN-2000; 2000WO-FR01566.
XX
XX 08-JUN-1999; 99FR-0007457.
XX
XX 06-AUG-1999; 99FR-0010378.
XX
XX (AVET ) AVENTIS PASTEUR.
XX
XX
XX Bachy M, Sodoyer R, Trannoy E;
XX
XX WPI; 2001-041317/05.
XX
XX
XX Oligonucleotides AAF28872-AAF28887 represent specific examples of an
XX immunostimulatory oligonucleotide of the invention which contains at
XX least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The
XX oligonucleotides do not contain any CG dinucleotides in which C are
XX unmethylated. The oligonucleotides are used as human immunostimulants
XX and as adjuvants in therapeutic and prophylactic vaccines for human use.
XX They induce proliferation of human lymphocytes, induce secretion of
XX cytokines, especially interleukin-10 or interferon-gamma and increase
XX expression of the CD86 activation marker or the CD25 cytokine receptor
XX on human B lymphocytes.
XX
XX Sequence 20 BP; 3 A; 1 C; 3 G; 13 T; 0 other;
SQ
Query Match 92.0%; Score 18.4; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 59;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTTATTAGTT 20
   ||||| ||||| ||||| |||||
Db 1 TTAGTTCTTACTTATTAGTT 20

RESULT 6
```

```
AAF28882
ID AAF28882 standard; DNA; 20 BP.
XX
AC AAF28882;
XX
DT 09-MAY-2001 (first entry)
XX
DE Immunostimulatory oligonucleotide #11 as vaccine adjuvant.
XX
KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KW phosphorothioate; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1..20
FT /tag= a
FT /note= "contain phosphorothioate internucleotide
FT linkages"
XX
PD WO200075304-A1.
XX
PN 14-DEC-2000.
XX
PF 08-JUN-2000; 2000WO-FR01566.
XX
PR 08-JUN-1999; 99FR-0007457.
PR 06-AUG-1999; 99FR-0010378.
XX
PA (AVER ) AVENTIS PASTEUR.
XX
PI Bachy M, Sodoyer R, Tranoy E;
XX
DR WPI; 2001-041317/05.
XX
PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
PT vaccines for human use, induce lymphocyte proliferation and cytokine
PT secretion -
XX
PS Example 4; Page 13; 30pp; French.
XX
CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an
CC immunostimulatory oligonucleotide of the invention which contains at
CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The
CC oligonucleotides do not contain any CG dinucleotides in which C are
CC unmethylated. The oligonucleotides are used as human immunostimulants
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
CC They induce proliferation of human lymphocytes, induce secretion of
CC cytokines, especially interleukin-10 or interferon-gamma and increase
CC expression of the CD86 activation marker or the CD25 cytokine receptor
CC on human B lymphocytes.
XX
SQ Sequence 20 BP; 3 A; 2 C; 3 G; 12 T; 0 other:
XX
Query Match 92.0%; Score 18.4; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 59;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTAGTCTTAGTTATTAGTT 20
DB 1 TTAGTCTTAGTTCTTAGTT 20
XX
RESULT 7
AAF28883
ID AAF28883 standard; DNA; 20 BP.
XX
AC AAF28883;
XX
DT 09-MAY-2001 (first entry)
XX
```

```
DE Immunostimulatory oligonucleotide #12 as vaccine adjuvant.
XX
KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KW phosphorothioate; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1..20
FT /tag= a
FT /note= "contain phosphorothioate internucleotide
FT linkages"
XX
PD WO200075304-A1.
XX
PN 14-DEC-2000.
XX
PF 08-JUN-2000; 2000WO-FR01566.
XX
PR 08-JUN-1999; 99FR-0007457.
PR 06-AUG-1999; 99FR-0010378.
XX
PA (AVER ) AVENTIS PASTEUR.
XX
PI Bachy M, Sodoyer R, Tranoy E;
XX
DR WPI; 2001-041317/05.
XX
PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
PT vaccines for human use, induce lymphocyte proliferation and cytokine
PT secretion -
XX
PS Example 4; Page 13; 30pp; French.
XX
CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an
CC immunostimulatory oligonucleotide of the invention which contains at
CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The
CC oligonucleotides do not contain any CG dinucleotides in which C are
CC unmethylated. The oligonucleotides are used as human immunostimulants
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
CC They induce proliferation of human lymphocytes, induce secretion of
CC cytokines, especially interleukin-10 or interferon-gamma and increase
CC expression of the CD86 activation marker or the CD25 cytokine receptor
CC on human B lymphocytes.
XX
SQ Sequence 20 BP; 3 A; 1 C; 4 G; 12 T; 0 other:
XX
Query Match 92.0%; Score 18.4; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 59;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTAGTCTTAGTTATTAGTT 20
DB 1 TTAGTCTTAGTTGTTAGTT 20
XX
RESULT 8
AAF28884
ID AAF28884 standard; DNA; 20 BP.
XX
AC AAF28884;
XX
DT 09-MAY-2001 (first entry)
XX
DE Immunostimulatory oligonucleotide #13 as vaccine adjuvant.
XX
KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KW phosphorothioate; ss.
XX
```

OS Synthetic.  
XX Key Location/Qualifiers  
FH modified\_base 1..20  
FT /\*tag= a  
FT /note= "contain phosphorothioate internucleotide  
FT linkages"  
XX  
XX WO200075304-A1.  
XX  
XX 14-DEC-2000.  
XX  
XX 08-JUN-2000; 2000WO-FR01566.  
XX  
XX 08-JUN-1999; 99FR-0007457.  
XX 06-AUG-1999; 99FR-0010378.  
XX  
XX (AVET ) AVENTIS PASTEUR.  
XX  
XX Bachy M, Sodoier R, Tranoy E;  
XX WPI; 2001-041317/05.  
XX  
XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in  
PT vaccines for human use, induce lymphocyte proliferation and cytokine  
PT secretion -  
XX  
XX Example 4; Page 13; 30pp; French.  
XX  
XX Oligonucleotides AAF28672-AAF28867 represent specific examples of an  
CC immunostimulatory oligonucleotide of the invention which contains at  
CC least one sequence 5'-TTNN2PT-3' where N1 and N2 are A, T, C or G. The  
CC oligonucleotides do not contain any CG dinucleotides in which C are  
CC unmethylated. The oligonucleotides are used as human immunostimulants  
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.  
CC They induce proliferation of human lymphocytes, induce secretion of  
CC cytokines, especially interleukin-10 or interferon-gamma and increase  
CC expression of the CD86 activation marker or the CD25 cytokine receptor  
CC on human B lymphocytes.  
XX  
XX Sequence 20 BP; 4 A; 0 C; 4 G; 12 T; 0 other;  
SQ  
Query Match 92.0%; Score 18.4; DB 22; Length 20;  
Best Local Similarity 95.0%; Pred. No. 59;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 TTAGTCTTACTTACTT 20  
Db 1 TTAGTCTTACTTACTT 20  
RESULT 9  
ABQ21728  
ID ABQ21728 standard; DNA; 749 BP.  
XX  
XX ABQ21728;  
XX  
XX 12-JUL-2002 (first entry)  
XX  
XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 8319.  
DE  
XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
KW drug; side effect; cancer; central nervous system; cardiovascular;  
KW gastrointestinal; respiratory system; single nucleotide polymorphism;  
KW SNP; cell differentiation; ds.  
XX  
XX Homo sapiens.  
XX  
XX OS  
XX PN WO200218632-A2.  
XX  
XX 07-MAR-2002.  
XX  
XX 01-SEP-2001; 2001WO-EPI0074.  
XX  
XX PF

XX  
XX 01-SEP-2000; 2000DE-1043826.  
PR  
XX 05-SEP-2000; 2000DE-1044543.  
XX  
XX  
XX (EPIG-) EPIGENOMICS AG.  
XX  
XX Olek A, Piepenbrock C, Berlin K, Guetig D;  
XX WPI; 2002-371829/40.  
XX  
XX  
XX Determining the degree of cytosine methylation in genomic DNA, useful  
PT for diagnosis and prognosis, comprises selective hybridization of  
PT amplicons from chemically treated DNA -  
XX  
XX Claim 12; 56pp + Sequence listing; 56pp; German.  
XX  
XX This invention describes a novel method for determining the degree of  
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
CC genomic sample of DNA. The sample is treated chemically to convert  
CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
CC DNA that contains the target C is amplified to form a labeled amplicon.  
CC The amplicon is hybridised to two classes, each with at least one  
CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers  
CC and the degree of hybridisation to both classes is determined from the  
CC label on the amplicon. From the ratio of labels hybridised to the two  
CC classes of oligomers, the degree of methylation is calculated. The method  
CC is used: (i) for diagnosis and/or prognosis of side effects of  
CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders  
CC of the central nervous, cardiovascular, gastrointestinal and respiratory  
CC systems etc., particularly by detecting mutations or single nucleotide  
CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue  
CC types and for investigating cell differentiation. The method allows the  
CC methylation status of many C residues to be determined simultaneously.  
CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the  
CC method for determining the degree of cytosine methylation described in  
CC the disclosure of the invention.  
XX  
XX Sequence 749 BP; 173 A; 54 C; 199 G; 322 T; 1 other;  
SQ  
Query Match 92.0%; Score 18.4; DB 24; Length 749;  
Best Local Similarity 95.0%; Pred. No. 52;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 TTAGTCTTACTTACTT 20  
Db 347 TTAGTCTTACTTACTT 366  
RESULT 10  
ABQ21729/c  
ID ABQ21729 standard; DNA; 749 BP.  
XX  
XX ABQ21729;  
XX  
XX 12-JUL-2002 (first entry)  
XX  
XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 8320.  
DE  
XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
KW drug; side effect; cancer; central nervous system; cardiovascular;  
KW gastrointestinal; respiratory system; single nucleotide polymorphism;  
KW SNP; cell differentiation; ds.  
XX  
XX OS  
XX PN WO200218632-A2.  
XX  
XX 07-MAR-2002.  
XX  
XX 01-SEP-2001; 2001WO-EPI0074.  
XX  
XX 01-SEP-2000; 2000DE-1043826.  
XX 05-SEP-2000; 2000DE-1044543.  
XX  
XX PR

```
XX (EPiG-) EPIGENOMICS AG.
PA Olek A, Piepenbrock C, Berlin K, Guetig D;
PI WPI; 2002-371829/40.
XX
DR Determining the degree of cytosine methylation in genomic DNA, useful
XX for diagnosis and prognosis, comprises selective hybridization of
PT amplicons from chemically treated DNA -
PS Claim 12: 56pp + Sequence Listing: 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one
CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
CC and the degree of hybridisation to both classes is determined from the
CC label on the amplicon. From the ratio of labels hybridised to the two
CC classes of oligomers, the degree of methylation is calculated. The method
CC is used: (i) for diagnosis and/or prognosis of side effects of
CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
CC of the central nervous, cardiovascular, gastrointestinal and respiratory
CC systems etc., particularly by detecting mutations or single nucleotide
CC polymorphisms (SNP's) and (ii) for differentiation of cell or tissue
CC types and for investigating cell differentiation. The method allows the
CC methylation status of many C residues to be determined simultaneously.
CC ABQ3410-ABQ54121 represent genomic DNA sequences used to illustrate the
CC method for determining the degree of cytosine methylation described in
CC the disclosure of the invention.
XX
SQ Sequence 749 BP; 322 A; 199 C; 54 G; 173 T; 1 other;
XX
Query Match 92.0%; Score 18.4; DB 24; Length 749;
Best Local Similarity 95.0%; Pred. No. 32;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTAGTCTTAGTTATTAGTT 20
DB 403 TTAGTTTCTTAGTTATTAGTT 384
XX
RESULT 11
AB132448
ID ABL32448 standard; DNA; 7238 BP.
XX
AC ABL32448;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 421.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antihaemic; cytosstatic; noctropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; arteriosclerosis; anaemia;
KW antinflammatory; cancer; eye disease; arteriosclerosis; AIDS;
KW acute myeloid leukaemia; Alzheimer's disease; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP07537.
XX
PR 30-JUN-2000; 2000DE-1032529.
XX
PS
```

```
PR 01-SEP-2000; 2000DE-1043826.
XX
XX (EPiG-) EPIGENOMICS AG.
PA Olek A, Piepenbrock C, Berlin K;
PI WPI; 2002-130909/17.
XX
DR Nucleic acid comprising fragment of chemically modified gene, useful
XX for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
PS Claim 1; SEQ ID NO 421; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/intestative bowel
CC diseases. The present sequence is a gene of the invention.
XX
SQ Sequence 7238 BP; 2414 A; 84 C; 1401 G; 3339 T; 0 other;
XX
Query Match 87.0%; Score 17.4; DB 24; Length 7238;
Best Local Similarity 94.7%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTAGTCTTAGTTATTAGT 19
DB 6286 TTAGTTTCTTAGTTATTAGT 6304
XX
RESULT 12
ABN80243
ID ABN80243 standard; DNA; 9507 BP.
XX
AC ABN80243;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human chemically modified disease associated gene SEQ ID NO 260.
XX
KW Human; development; homeobox gene; HOX; diabetes; cancer; apoptosis;
KW heart disease; epilepsy; epistone deacetylation; muscular dystrophy;
KW dwarfism; single nucleotide polymorphism; SNP; cytosine methylation;
KW antidiabetic; cytosstatic; anticonvulsant; ds.
XX
OS Homo sapiens.
XX
OS Synthetic.
XX
PN WO200200927-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP07536.
XX
PR 30-JUN-2000; 2000DE-1032529.
XX
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPiG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130908/17.
XX
PD Novel nucleic acid useful for diagnosis and therapy of diseases
XX associated with development genes such as diabetes, comprises a
XX sequence of a segment of chemically pretreated DNA of genes associated
XX with development -
PS Claim 1; SEQ ID NO 260; 27pp; English.
```



XX The invention relates to a nucleic acid (I) comprising a sequence at  
CC least 18 bases in length of a segment of chemically pretreated DNA (II)  
CC of genes associated with development selected from 87 genes listed in  
CC the specification such as ACCPN, ADFN, or AFDI and comprising one of 350  
CC sequences (ABN79984-ABN80333) or their complements. The invention is  
CC useful for the diagnosis or therapy of diseases associated with  
CC development genes, in particular disease related to homeobox containing  
CC genes (HOX), like diabetes, cancer, apoptosis related diseases, syndromes  
CC associated with congenital heart disease, epilepsy, diseases related to  
CC histone deacetylation, Currarino syndrome, diseases related with the  
CC development of the brain and limb girdle muscular dystrophy and dwarfism.  
CC Oligomers specific to each of the genes are useful for detecting the  
CC methylation state of all CpG dinucleotides within the 350 sequences or  
CC (II) and their complementary sequences, as primer oligonucleotides for  
CC the amplification of the 350 sequences, (II) and/or their complements and  
CC as oligomer probes for detecting the cytosine methylation state and/or  
CC single nucleotide polymorphisms (SNPs).  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification but is based on sequence information supplied to Derwent by  
CC the European Patent Office.

SO Sequence 9507 BP; 2378 A; 326 C; 2397 G; 4406 T; 0 other;

Query Match 87.0%; Score 17.4; DB 24; Length 9507;  
Best Local Similarity 94.7%; Pred. No. 1.3e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTTATTACT 19  
||||| |||||||||  
DB 106 TTAGTCTTACTTATTACT 124

## RESULT 13

AAE28873  
ID AAE28873 standard; DNA; 20 BP.

AC AAE28873;

DT 09-MAY-2001 (first entry)

DE Immunostimulatory oligonucleotide #2 as vaccine adjuvant.

KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;  
KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;  
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;  
KW phosphorothioate; ss.

KM  
XX Synthetic.

OS  
XX

Key Location/Qualifiers

FT modified\_base 1..20

FT /tag= a  
FT /note= "contain phosphorothioate internucleotide  
FT linkages"

FT WO200075304-A1.

PN 14-DEC-2000.

PD 08-JUN-2000; 2000WO-FR01566.

PR 08-JUN-1999; 99FR-0007457.

PR 06-AUG-1999; 99FR-0010378.

XX (AVET ) AVENTIS PASTEUR.

XX Bachy M, Sodoyer R, Tranoy E;

XX WPI: 2001-041317/05.

XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in  
PT vaccines for human use, induce lymphocyte proliferation and cytokine

PT secretion -  
XX  
PS Example 4; Page 13; 30pp; French.

CC Oligonucleotides AAE28872-AAE28887 represent specific examples of an  
CC immunostimulatory oligonucleotide of the invention which contains at  
CC least one sequence 5'-TTNN2TT-3' where N1 and N2 are A, T, C or G. The  
CC oligonucleotides do not contain any CG dinucleotides in which C are  
CC unethyalted. The oligonucleotides are used as human immunostimulants  
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.  
CC They induce proliferation of human lymphocytes, induce secretion of  
CC cytokines, especially interleukin-10 or interferon-gamma and increase  
CC expression of the CD86 activation marker or the CD25 cytokine receptor  
CC on human B lymphocytes.

SO Sequence 20 BP; 4 A; 0 C; 3 G; 13 T; 0 other;

Query Match 84.0%; Score 16.8; DB 22; Length 20;  
Best Local Similarity 90.0%; Pred. No. 2.8e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTTATTACTT 20  
||||| ||||||| |||||  
DB 1 TTAGTCTTACTTATTACTT 20

## RESULT 14

AAE28874  
ID AAE28874 standard; DNA; 20 BP.

AC AAE28874;

DT 09-MAY-2001 (first entry)

DE Immunostimulatory oligonucleotide #3 as vaccine adjuvant.

KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;  
KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;  
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;  
KW phosphorothioate; ss.

KM  
XX Synthetic.

OS  
XX

Key Location/Qualifiers

FT modified\_base 1..20

FT /tag= a  
FT /note= "contain phosphorothioate internucleotide  
FT linkages"

FT WO200075304-A1.

PN 14-DEC-2000.

PD 08-JUN-2000; 2000WO-FR01566.

PR 08-JUN-1999; 99FR-0007457.

PR 06-AUG-1999; 99FR-0010378.

XX (AVET ) AVENTIS PASTEUR.

XX Bachy M, Sodoyer R, Tranoy E;

XX WPI: 2001-041317/05.

XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in  
PT vaccines for human use, induce lymphocyte proliferation and cytokine  
PT secretion -

PS Example 4; Page 13; 30pp; French.

CC Oligonucleotides AAE28872-AAE28887 represent specific examples of an  
CC immunostimulatory oligonucleotide of the invention which contains at  
CC least one sequence 5'-TTNN2TT-3' where N1 and N2 are A, T, C or G. The

CC oligonucleotides do not contain any CG dinucleotides in which C are  
CC unmethylated. The oligonucleotides are used as human immunostimulants  
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.  
CC They induce proliferation of human lymphocytes, induce secretion of  
CC cytokines, especially interleukin-10 or interferon-gamma and increase  
CC expression of the CD86 activation marker or the CD25 cytokine receptor  
CC on human B lymphocytes.  
XX  
SQ Sequence 20 BP; 4 A; 1 C; 3 G; 12 T; 0 other;  
Query Match 84.0%; Score 16.8; DB 22; Length 20;  
Best Local Similarity 90.0%; Pred. No. 2.8e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 TTAGTCTTAGTTATTAGTT 20  
||||| ||||| |||||  
Db 1 TTAGTTATTAGTTCTTAGTT 20  
RESULT 15  
AAF28875  
ID AAF28875 standard; DNA: 20 BP.  
XX  
AC AAF28875;  
XX  
DT 09-MAY-2001 (first entry)  
XX  
DE Immunostimulatory oligonucleotide #4 as vaccine adjuvant.  
XX  
KM Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;  
KM prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;  
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;  
KW phosphorothioate; ss.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT modified\_base 1..20  
FT /tag- a  
FT /note- "contain phosphorothioate Internucleotide  
linkages"  
XX  
PN WO200075304-A1.  
XX  
PD 14-DEC-2000.  
XX  
PF 08-JUN-2000; 2000WO-FR01566.  
XX  
PR 08-JUN-1999; 99FR-0007457.  
PR 06-AUG-1999; 99FR-0010378.  
XX  
PA (AVET ) AVENTIS PASTEUR.  
XX  
PI Bachy M, Sodey R, Tranoy E;  
XX  
DR WPI; 2001-041317/05.  
XX  
PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in  
PT vaccines for human use, induce lymphocyte proliferation and cytokine  
PT secretion -  
XX  
PS Example 4; Page 13; 30pp; French.  
XX  
CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an  
CC immunostimulatory oligonucleotide of the invention which contains at  
CC least one sequence 5'-TTNIN2YT-3' where N1 and N2 are A, T, C or G. The  
CC oligonucleotides do not contain any CG dinucleotides in which C are  
CC unmethylated. The oligonucleotides are used as human immunostimulants  
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.  
CC They induce proliferation of human lymphocytes, induce secretion of  
CC cytokines, especially interleukin-10 or interferon-gamma and increase  
CC expression of the CD86 activation marker or the CD25 cytokine receptor  
CC on human B lymphocytes.

XX  
SQ Sequence 20 BP; 4 A; 0 C; 4 G; 12 T; 0 other;  
Query Match 84.0%; Score 16.8; DB 22; Length 20;  
Best Local Similarity 90.0%; Pred. No. 2.8e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 TTAGTCTTAGTTATTAGTT 20  
||||| ||||| |||||  
Db 1 TTAGTTATTAGTTCTTAGTT 20  
Search completed: April 6, 2003, 12:05:22  
Job time : 152 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 11:12:37 ; Search time 45.25 Seconds  
(without alignments)  
135.548 Million cell updates/sec

Title: US-09-980-265-9

Perfect score: 1 ttactcttagtattagtt 20

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA: \*  
1: /cgn2-6/ptodata/2/lna/5A\_COMB.seq: \*  
2: /cgn2-6/ptodata/2/lna/5B\_COMB.seq: \*  
3: /cgn2-6/ptodata/2/lna/6A\_COMB.seq: \*  
4: /cgn2-6/ptodata/2/lna/6B\_COMB.seq: \*  
5: /cgn2-6/ptodata/2/lna/PTCUTS\_COMB.seq: \*  
6: /cgn2-6/ptodata/2/lna/Backfills1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16.8	84.0	59065	4 US-09-813-817-3	Sequence 3, Appli
2	16.8	84.0	59065	4 US-09-978-197-3	Sequence 3, Appli
3	15.8	79.0	29793	4 US-09-302-812-38	Sequence 38, Appli
4	15.8	79.0	29793	4 US-09-511-477-38	Sequence 38, Appli
5	15.8	79.0	29793	4 US-09-511-507-38	Sequence 38, Appli
6	15.2	76.0	403	2 US-08-565-166-30	Sequence 30, Appli
7	15.2	76.0	1203	3 US-09-122-443-3	Sequence 3, Appli
8	15.2	76.0	1578	4 US-09-134-001C-677	Sequence 677, App
9	15.2	76.0	1602	2 US-08-770-544-3	Sequence 3, Appli
10	15.2	76.0	1857	4 US-09-299-378-3	Sequence 3, Appli
11	15.2	76.0	2823	1 US-08-398-008A-1	Sequence 1, Appli
12	15.2	76.0	2823	2 US-08-893-333-1	Sequence 1, Appli
13	15.2	76.0	5049	1 US-08-336-345-2	Sequence 2, Appli
14	15.2	76.0	5049	2 US-08-647-655-2	Sequence 2, Appli
15	15.2	76.0	5935	4 US-09-178-973B-17	Sequence 17, Appli
16	15.2	76.0	5935	4 US-09-419-568F-29	Sequence 29, Appli
17	15.2	76.0	5935	4 US-09-354-243B-29	Sequence 29, Appli
18	14.8	74.0	20	1 US-08-588-821-76	Sequence 76, Appli
19	14.8	74.0	20	1 US-08-915-214-76	Sequence 76, Appli
20	14.8	74.0	20	2 US-09-005-532-76	Sequence 76, Appli
21	14.8	74.0	200	2 US-08-875-972-27	Sequence 27, Appli
22	14.8	74.0	292	1 US-08-644-664B-41	Sequence 41, Appli
23	14.8	74.0	292	2 US-08-761-227A-41	Sequence 41, Appli
24	14.8	74.0	348	1 US-08-220-606B-54	Sequence 54, Appli
25	14.8	74.0	437	1 US-09-397-787-315	Sequence 315, App
26	14.8	74.0	463	4 US-09-385-982-407	Sequence 407, App
27	14.8	74.0	626	4 US-09-328-111-519	Sequence 519, App

C	28	14.8	74.0	627	4 US-09-385-982-186	Sequence 186, App
	29	14.8	74.0	2153	6 5225348-2	Patent No. 5225348
	30	14.8	74.0	1288	2 US-08-371-377-16	Sequence 16, Appl
	31	14.8	74.0	3073	1 US-07-688-352C-31	Sequence 31, Appl
	32	14.8	74.0	3073	2 US-08-474-379C-31	Sequence 31, Appl
	33	14.8	74.0	3073	3 US-09-146-249A-31	Sequence 31, Appl
	34	14.8	74.0	3073	4 US-08-206-188B-31	Sequence 31, Appl
	35	14.8	74.0	3073	5 PCP-US91-02714-30	Sequence 30, Appl
	36	14.8	74.0	4695	4 US-09-309-572-9	Sequence 9, Appli
	37	14.8	74.0	4695	6 5225348-3	Patent No. 5225348
	38	14.8	74.0	11282	4 US-09-754-250-3	Sequence 3, Appli
	39	14.8	74.0	168575	4 US-09-426-290-1	Sequence 1, Appli
	40	14.8	74.0	4411529	4 US-09-103-840A-1	Sequence 1, Appli
	41	14.4	72.0	5010	3 US-09-161-244-1	Sequence 442, App
	42	14.2	71.0	362	4 US-09-605-785-442	Sequence 442, App
	43	14.2	71.0	362	4 US-09-439-313-442	Sequence 442, App
	44	14.2	71.0	362	4 US-09-352-616A-442	Sequence 46, Appl
	45	14.2	71.0	427	4 US-09-556-877-46	

#### ALIGNMENTS

```
RESULT 1
US-09-813-817-3
; Sequence 3, Application US/09813817
; Patent No. 6340583
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: C1001178
; CURRENT APPLICATION NUMBER: US/09/813,817
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 59065
; TYPE: DNA
; ORGANISM: Human
US-09-813-817-3

Query Match      84.0%  Score 16.8;  DB 4;  Length 59065;
Best Local Similarity 90.0%  Pred. No. 33;
Matches 18;  Conservative 0;  Mismatches 2;  Indels 0;  Gaps 0;

QY      1  TTAGTCTTACTGTTATTTAGTT 20
        || ||||| ||||| |||||
DB 15314 TTGTCTTGTGTTATTTAGTT 15333

RESULT 2
US-09-978-197-3
; Sequence 3, Application US/09978197
; Patent No. 6403353
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: C1001178DIV
; CURRENT APPLICATION NUMBER: US/09/978,197
; CURRENT FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 09/813,817
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 59065
; TYPE: DNA
; ORGANISM: Human
US-09-978-197-3
```

Query Match 84.0%; Score 16.8; DB 4; Length 59065;  
Best Local Similarity 90.0%; Pred. No. 32;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAGTCTTACTTACTT 20  
Db 15314 TTGTCTTCTTGTATTAATT 15333

## RESULT 3

US-09-302-812-38  
; Sequence 38, Application US/09302812B  
; Patent No. 6335148  
; GENERAL INFORMATION:  
; APPLICANT: JACOBSON, Myron K.  
; APPLICANT: JACOBSON, Elaine L.  
; APPLICANT: AM, Jean-Christophe  
; APPLICANT: LIN, Winston  
; TITLE OF INVENTION: GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOHYDROLASE (PARG) ENZ  
; TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACTIVE  
; FILE REFERENCE: NIAD 201  
; CURRENT APPLICATION NUMBER: US/09/302,812B  
; CURRENT FILING DATE: 1999-04-30  
; EARLIER APPLICATION NUMBER: 60/083,768  
; PRIOR FILING DATE: 1998-05-01  
; NUMBER OF SEQ ID NOS: 38  
; SEQ ID NO 38  
; LENGTH: 29793  
; TYPE: DNA  
; ORGANISM: Caenorhabditis elegans  
; FEATURE:  
US-09-302-812-38

Query Match 79.0%; Score 15.8; DB 4; Length 29793;  
Best Local Similarity 89.5%; Pred. No. 86;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TAGTCTTACTTACTT 20  
Db 20450 TAGCTCTTACTTACTTATT 20468

## RESULT 4

US-09-511-477-38  
; Sequence 38, Application US/09511477  
; Patent No. 6337202  
; GENERAL INFORMATION:  
; APPLICANT: JACOBSON, Myron K.  
; APPLICANT: JACOBSON, Elaine L.  
; APPLICANT: AME, Jean-Christophe  
; APPLICANT: LIN, Winston  
; TITLE OF INVENTION: GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOHYDROLASE (PARG) EN  
; TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACTIV  
; TITLE OF INVENTION: THEREWITH  
; FILE REFERENCE: NIAD 201  
; CURRENT APPLICATION NUMBER: US/09/511,477  
; CURRENT FILING DATE: 2000-02-23  
; PRIOR APPLICATION NUMBER: 09/302,812  
; PRIOR FILING DATE: 1999-04-30  
; NUMBER OF SEQ ID NOS: 38  
; SEQ ID NO 38  
; LENGTH: 29793  
; TYPE: DNA  
; ORGANISM: Caenorhabditis elegans  
; FEATURE:  
US-09-511-477-38

Query Match 79.0%; Score 15.8; DB 4; Length 29793;  
Best Local Similarity 89.5%; Pred. No. 86;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TAGTCTTACTTACTT 20  
Db 20450 TAGCTCTTACTTACTTATT 20468

## RESULT 5

US-09-511-507-38  
; Sequence 38, Application US/09511507  
; Patent No. 6395543  
; GENERAL INFORMATION:  
; APPLICANT: JACOBSON, Myron K.  
; APPLICANT: JACOBSON, Elaine L.  
; APPLICANT: AME, Jean-Christophe  
; APPLICANT: LIN, Winston  
; TITLE OF INVENTION: GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOHYDROLASE (PARG)  
; TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREAC  
; FILE REFERENCE: NIAD 201  
; CURRENT APPLICATION NUMBER: US/09/511,507  
; CURRENT FILING DATE: 2000-02-23  
; PRIOR APPLICATION NUMBER: 09/302,812  
; PRIOR FILING DATE: 1999-04-30  
; NUMBER OF SEQ ID NOS: 38  
; SEQ ID NO 38  
; LENGTH: 29793  
; TYPE: DNA  
; ORGANISM: Caenorhabditis elegans  
; FEATURE:  
US-09-511-507-38

Query Match 79.0%; Score 15.8; DB 4; Length 29793;  
Best Local Similarity 89.5%; Pred. No. 86;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TAGTCTTACTTACTT 20  
Db 20450 TAGCTCTTACTTACTTATT 20468

## RESULT 6

US-08-569-166-30  
; Sequence 30, Application US/08569166  
; Patent No. 5830722  
; GENERAL INFORMATION:  
; APPLICANT: NICOLAS, LUC  
; APPLICANT: CHARLES, JEAN-FRANCOIS  
; APPLICANT: DELECLUSE, AMELLE  
; APPLICANT: BARLOY, FREDERIQUE  
; TITLE OF INVENTION: CLOSTRIDIUM BIFERMENTANS DNA FRAGMENT  
; TITLE OF INVENTION: BEARING GENES CODING FOR PROTEINS LINKED TO AN  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/569,166  
; FILING DATE: 05-JUL-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR94/00768  
; FILING DATE: 24-JUN-1994  
; PRIOR APPLICATION DATA:

```

? APPLICATION NUMBER: FR 93/07795
? FILING DATE: 25-JUN-1993
? ATTORNEY/AGENT INFORMATION:
? NAME: OBLON, NORMAN F.
? REGISTRATION NUMBER: 24,618
? REFERENCE/DOCKET NUMBER: 660-106-0 PCT
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 703-413-3000
? TELEFAX: 703-413-2220
? INFORMATION FOR SEQ ID NO: 30:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 403 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? US-08-569-166-30

Query Match
Best Match Similarity 76.0%; Score 15.2; DB 2; Length 403;
Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTATTAGTT 20
Db 99 TTACATCTTGTATTAGTT 118

RESULT 7
US-09-122-443-3/C
? Sequence 3, Application US/09122443
? Patent No. 6060284
? GENERAL INFORMATION:
? APPLICANT: Bazan, J. Fernando
? TITLE OF INVENTION: MAMMALIAN CYTOKINE; RELATED REAGENTS
? NUMBER OF SEQUENCES: 16
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: DNAX Research Institute
? STREET: 901 California Avenue
? CITY: Palo Alto
? STATE: California
? COUNTRY: USA
? ZIP: 94304-1104
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/122,443
? FILING DATE: 24-JUL-1998
? CLASSIFICATION: 436
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/053,765
? FILING DATE: 25-JUL-1997
? ATTORNEY/AGENT INFORMATION:
? NAME: Ching, Edwin P.
? REGISTRATION NUMBER: 34,090
? REFERENCE/DOCKET NUMBER: DX0758K1
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (650)852-9196
? TELEFAX: (650)496-1200
? INFORMATION FOR SEQ ID NO: 3:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1203 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: CDNA
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 113..700
? FEATURE:
? NAME/KEY: mat_peptide
```

```

? LOCATION: 176..700
? US-09-122-443-3

Query Match
Best Match Similarity 76.0%; Score 15.2; DB 3; Length 1203;
Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTATTAGTT 20
Db 1177 TTACTTCTAGTACTAGAT 1158

RESULT 8
US-09-134-001C-677
? Sequence 677, Application US/09134001C
? Patent No. 6380370
? GENERAL INFORMATION:
? APPLICANT: Lynn Doucette-Stamm et al
? TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
? TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
? FILE REFERENCE: GTC-007
? CURRENT APPLICATION NUMBER: US/09/134,001C
? CURRENT FILING DATE: 1998-08-13
? PRIOR APPLICATION NUMBER: US 60/064,964
? PRIOR FILING DATE: 1997-11-08
? PRIOR APPLICATION NUMBER: US 60/055,779
? PRIOR FILING DATE: 1997-08-14
? NUMBER OF SEQ ID NOS: 5674
? SEQ ID NO 677
? LENGTH: 1578
? TYPE: DNA
? ORGANISM: Staphylococcus epidermidis
? US-09-134-001C-677

Query Match
Best Match Similarity 76.0%; Score 15.2; DB 4; Length 1578;
Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTATTAGTT 20
Db 796 TTTGTATTAGTTTGT 815

RESULT 9
US-08-770-544-3
? Sequence 3, Application US/08770544
? Patent No. 5907085
? GENERAL INFORMATION:
? APPLICANT: Gonsalves, Dennis
? APPLICANT: Ling, Kai-Shu
? TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS PROTEINS AND
? TITLE OF INVENTION: THEIR USES
? NUMBER OF SEQUENCES: 32
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
? STREET: Clinton Square, P.O. Box 1051
? CITY: Rochester
? STATE: New York
? COUNTRY: U.S.A.
? ZIP: 14603
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/770,544
? FILING DATE:
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60009008
? FILING DATE: 21-DEC-1995
? ATTORNEY/AGENT INFORMATION:
```

```

; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/621
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1304
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1602 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-770-344-3

Query Match      76.0%; Score 15.2; DB 2; Length 1602;
Best Local Similarity 85.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 TTAGTCTTAGTATTAGTT 20
        ||||||| ||||| |||||
DB      1190 TTGTCTTAGTATTAGTT 1209

RESULT 10
US-09-299-378-3
; Sequence 3, Application US/09299378
; Patent No. 6245545
; GENERAL INFORMATION:
; APPLICANT: Higgins, Lauren S.
; APPLICANT: Dalton, Michael A.
; APPLICANT: Kong, HuiMin
; TITLE OF INVENTION: Method For Cloning And Producing The Swai Restriction
; FILE REFERENCE: Swai
; CURRENT APPLICATION NUMBER: US/09/299,378
; CURRENT FILING DATE: 1999-04-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1857
; TYPE: DNA
; ORGANISM: Staphylococcus warneri
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1854)
US-09-299-378-3

Query Match      76.0%; Score 15.2; DB 4; Length 1857;
Best Local Similarity 85.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 TTAGTCTTAGTATTAGTT 20
        ||| ||||| ||||| ||||| ||
DB      1703 TTAATCTTAGTATTATTT 1722

RESULT 11
US-08-398-008A-1
; Sequence 1, Application US/08398008A
; Patent No. 5665588
; GENERAL INFORMATION:
; APPLICANT: Kornbluth, Jacki
; TITLE OF INVENTION: DNA Encoding Natural Killer Lytic Associated
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gildbreth & Adler, P.C.
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77071
```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB storage
; COMPUTER: MACINTOSH IICx
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/398,008A
; FILING DATE: March 2, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/126,501
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler, Dr. Benjamin Aaron
; REGISTRATION NUMBER: 35,423
; REFERENCE/DOCKET NUMBER: D5705CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 777-2321
; TELEFAX: (713) 777-6908
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2823 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double-stranded
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: no
; ANTI-SENSE: no
US-08-398-008A-1

Query Match      76.0%; Score 15.2; DB 1; Length 2823;
Best Local Similarity 85.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 TTAGTCTTAGTATTAGTT 20
        ||||||| ||||| ||||| |||||
DB      2654 TTAGTTTAATTATTAGTT 2673

RESULT 12
US-08-893-333-1
; Sequence 1, Application US/08893333A
; Patent No. 5981705
; GENERAL INFORMATION:
; APPLICANT: Kornbluth, Jacki
; TITLE OF INVENTION: DNA Encoding Natural Killer Lytic Associated Protein
; FILE REFERENCE: D5705CIP/D
; CURRENT APPLICATION NUMBER: US/08/893,333A
; CURRENT FILING DATE: 1997-07-16
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 1
; LENGTH: 2823
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: 190..1953
; OTHER INFORMATION: CDS
US-08-893-333-1

Query Match      76.0%; Score 15.2; DB 2; Length 2823;
Best Local Similarity 85.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 TTAGTCTTAGTATTAGTT 20
        ||||||| ||||| ||||| |||||
DB      2654 TTAGTTTAATTATTAGTT 2673

RESULT 13
US-08-336-345-2
; Sequence 2, Application US/08336345
; Patent No. 5614510
```

```

: GENERAL INFORMATION:
: APPLICANT: Parrish, Colin R.
: APPLICANT: Gruenberg, Allen
: APPLICANT: Carmichael, Ireland E.
: TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: NY
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/336,345
: FILING DATE:
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Gordon, Jennifer
: REGISTRATION NUMBER: 30753
: REFERENCE/DOCKET NUMBER: 7937-006
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-8864/9741
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5049 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: unknown
: MOLECULE TYPE: DNA (genomic)
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Parvovirus
:
: US-08-336-345-2
:
: Query Match 76.0%; Score 15.2; DB 1; Length 5049;
: Best Local Similarity 85.0%; Pred. No. 1.6e+02;
: Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
:
: Oy 1 TTAGTCTTAGTATTAGTT 20
: ||||| || ||||| |||||
: Db 4745 TTAGATCATGTTAGTTAGTT 4764
:
: RESULT 14
: US-08-647-655-2
: Sequence 2, Application US/08647655
: Patent No. 5885585
: GENERAL INFORMATION:
: APPLICANT: Parrish, Colin R.
: APPLICANT: Gruenberg, Allen
: APPLICANT: Carmichael, Ireland E.
: TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/647,655

```

```

: FILING DATE: On Even Date Herewith
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Gordon, Jennifer
: REGISTRATION NUMBER: 30,753
: REFERENCE/DOCKET NUMBER: 7937-008
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-8864/9741
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5049 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: unknown
: MOLECULE TYPE: DNA (genomic)
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Parvovirus
:
: US-08-647-655-2
:
: Query Match 76.0%; Score 15.2; DB 2; Length 5049;
: Best Local Similarity 85.0%; Pred. No. 1.6e+02;
: Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
:
: Oy 1 TTAGTCTTAGTATTAGTT 20
: ||||| || ||||| |||||
: Db 4745 TTAGATCATGTTAGTTAGTT 4764
:
: RESULT 15
: US-09-178-973B-17
: Sequence 17, Application US/09178973B
: Patent No. 6274710
: GENERAL INFORMATION:
: APPLICANT: Dumoulier, Laure
: APPLICANT: Louhed, Jamila
: APPLICANT: Renaud, Jean-Christophe
: TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible
: TITLE OF INVENTION: (TIRs)
: FILE REFERENCE: LOD 5543
: CURRENT APPLICATION NUMBER: US/09/178,973B
: CURRENT FILING DATE: 1998-10-26
: NUMBER OF SEQ ID NOS: 17
: SEQ ID NO 17
: LENGTH: 5935
: TYPE: DNA
: ORGANISM: Mus musculus
:
: US-09-178-973B-17
:
: Query Match 76.0%; Score 15.2; DB 4; Length 5935;
: Best Local Similarity 85.0%; Pred. No. 1.6e+02;
: Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
:
: Oy 1 TTAGTCTTAGTATTAGTT 20
: ||||| || ||||| |||||
: Db 2996 TCATTCTTAGTATTATT 3015
:
: Search completed: April 6, 2003, 11:55:37
: Job time : 79.25 secs

```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 11:24:32 ; Search time 58 Seconds  
(without alignments)  
302.472 Million cell updates/sec

Title: US-09-980-265-9  
Perfect score: 20  
Sequence: 1 ttacttcttagttatgatt 20

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 593429 seqs, 438583890 residues  
Total number of hits satisfying chosen parameters: 1186858

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

Published\_Applications\_NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCr\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCrUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	16.8	84.0	548	10 US-09-924-035A-396	Sequence 396, App
2	16.8	80.0	2000	9 US-09-938-842A-4669	Sequence 4669, App
3	15.8	79.0	401	9 US-09-946-807-941	Sequence 941, App
4	15.8	79.0	401	10 US-09-795-668-941	Sequence 941, App
5	15.8	79.0	401	10 US-09-795-668-941	Sequence 941, App
C 6	15.8	79.0	421	9 US-09-954-551-424	Sequence 424, App
C 7	15.8	79.0	421	10 US-09-880-107-220	Sequence 220, App
8	15.8	79.0	492	9 US-10-091-504-1359	Sequence 1359, App
9	15.8	79.0	492	9 US-10-091-504-1360	Sequence 1360, App
10	15.8	79.0	492	9 US-10-091-504-1361	Sequence 1361, App
11	15.8	79.0	492	10 US-09-764-869-1359	Sequence 1359, App
12	15.8	79.0	492	10 US-09-764-869-1360	Sequence 1360, App
13	15.8	79.0	492	10 US-09-764-869-1361	Sequence 1361, App
C 14	15.8	79.0	551	7 US-08-781-986A-994	Sequence 994, App
C 15	15.8	79.0	573	9 US-09-232-785-330	Sequence 330, App
16	15.8	79.0	657	9 US-09-938-842A-3055	Sequence 3055, App
17	15.8	79.0	1305	10 US-09-887-576-491	Sequence 491, App
18	15.8	79.0	1305	10 US-09-887-576-491	Sequence 491, App
19	15.8	79.0	2000	9 US-09-938-842A-4454	Sequence 4454, App

20	15.8	79.0	2000	10 US-09-887-576-97	Sequence 97, App1
21	15.8	79.0	2002	10 US-09-887-576-3	Sequence 3, App1
22	15.8	79.0	2010	10 US-09-887-576-2	Sequence 2, App1
C 23	15.8	79.0	8253	9 US-10-091-572-622	Sequence 622, App
C 24	15.8	79.0	13392	9 US-10-091-572-621	Sequence 621, App
25	15.8	79.0	29793	10 US-09-973-451-38	Sequence 38, App1
C 26	15.4	77.0	1087	9 US-09-822-846-475	Sequence 475, App
27	15.4	77.0	659158	9 US-09-771-208-20	Sequence 20, App1
28	15.2	76.0	407	10 US-09-878-574-3626	Sequence 3626, App
29	15.2	76.0	572	10 US-09-864-761-12035	Sequence 12035, App
30	15.2	76.0	658	9 US-09-938-842A-4830	Sequence 4830, App
C 31	15.2	76.0	1557	9 US-10-001-873-1	Sequence 1, App1
32	15.2	76.0	1656	10 US-09-070-927A-499	Sequence 499, App
C 33	15.2	76.0	1684	9 US-09-938-842A-4070	Sequence 4070, App
C 34	15.2	76.0	2049	7 US-08-781-986A-291	Sequence 291, App
C 35	15.2	76.0	3283	10 US-09-925-302-144	Sequence 144, App
C 36	15.2	76.0	3772	10 US-09-822-830A-570	Sequence 570, App
C 37	15.2	76.0	4920	7 US-08-781-986A-26	Sequence 26, App1
38	15.2	76.0	5150	9 US-09-764-904-94	Sequence 94, App1
39	15.2	76.0	5150	9 US-10-091-548-94	Sequence 94, App1
40	15.2	76.0	5150	10 US-09-764-860-1138	Sequence 1138, App
41	15.2	76.0	5935	10 US-09-751-797-29	Sequence 29, App1
42	15.2	76.0	32174	9 US-09-860-670-232	Sequence 232, App
43	15.2	76.0	32174	9 US-09-764-904-90	Sequence 90, App1
44	15.2	76.0	32174	9 US-10-091-548-90	Sequence 90, App1
45	15.2	76.0	32174	10 US-09-908-711-158	Sequence 158, App

## ALIGNMENTS

RESULT 1  
US-09-924-035A-396/c  
Sequence 396, Application US/09924035A  
Patent No. US20020142319A1  
GENERAL INFORMATION:  
APPLICANT: Glaxo, Jrn  
TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
FILE REFERENCE: 2011US  
CURRENT FILING DATE: 2000-08-11  
CURRENT APPLICATION NUMBER: US/09/924,035A  
PRIOR APPLICATION NUMBER: US 60/148,784  
PRIOR FILING DATE: 1999-08-13  
NUMBER OF SEQ ID NOS: 900  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 396  
LENGTH: 548  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
NAME/KEY: misc-feature  
LOCATION: (1)...(548)  
OTHER INFORMATION: n = A,T,C or G  
US-09-924-035A-396  
Query Match  
Best Local Similarity 84.0% Score 16.8; DB 10; Length 548;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Oy 1 TTACTTCTTACTTACTT 20  
DB 172 TTACTTCTTACTTACTT 153  
RESULT 2  
US-09-938-842A-4669  
Sequence 4669, Application US/09938842A  
Patent No. US20020160378A1  
GENERAL INFORMATION:  
APPLICANT: Harper, Jeff  
APPLICANT: Kreps, Joel  
APPLICANT: Wang, Xun

```

; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 4669
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-4669

```

```

Query Match      80.0%; Score 16; DB 9; Length 2000;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TTAGTCTTAGTTATT 16
Db 1304 TTAGTCTTAGTTATT 1319

```

```

RESULT 3
US-09-946-807-941
; Sequence 941, Application US/09946807
; Patent No. US20020165144A1
; GENERAL INFORMATION:
; APPLICANT: Stefansson, Hreinn
; APPLICANT: Steinthorsdottir, Valgerdur
; APPLICANT: Gulcher, Jeffrey R.
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
; FILE REFERENCE: 2345.2004-001
; CURRENT APPLICATION NUMBER: US/09/946,807
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US/09/795,668
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/515,716
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 1531
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 941
; LENGTH: 401
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-946-807-941

```

```

Query Match      79.0%; Score 15.8; DB 9; Length 401;
Best Local Similarity 89.5%; Pred. No. 5.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 TTAGTCTTAGTTATT 19
Db 254 TTAGTCTTAGTTATTAGT 272

```

```

RESULT 4
US-09-795-668-941
; Sequence 941, Application US/09795668
; Patent No. US20020045577A1
; GENERAL INFORMATION:
; APPLICANT: Stefansson, Hreinn
; APPLICANT: Steinthorsdottir, Valgerdur
; APPLICANT: Gulcher, Jeffrey R.
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
; FILE REFERENCE: 2345.2004-001
; CURRENT APPLICATION NUMBER: US/09/795,668
; CURRENT FILING DATE: 2001-02-28

```

```

; PRIOR APPLICATION NUMBER: US 09/515,716
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 1531
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 941
; LENGTH: 401
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-795-668-941

```

```

Query Match      79.0%; Score 15.8; DB 10; Length 401;
Best Local Similarity 89.5%; Pred. No. 5.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 TTAGTCTTAGTTATTAGT 19
Db 254 TTAGTCTTAGTTATTAGT 272

```

```

RESULT 5
US-09-795-686-941
; Sequence 941, Application US/09795686
; Patent No. US20020094954A1
; GENERAL INFORMATION:
; APPLICANT: Stefansson, Hreinn
; APPLICANT: Steinthorsdottir, Valgerdur
; APPLICANT: Gulcher, Jeffrey R.
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
; FILE REFERENCE: 2345.2005-001
; CURRENT APPLICATION NUMBER: US/09/795,686
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/515,715
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 1531
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 941
; LENGTH: 401
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-795-686-941

```

```

Query Match      79.0%; Score 15.8; DB 10; Length 401;
Best Local Similarity 89.5%; Pred. No. 5.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 TTAGTCTTAGTTATTAGT 19
Db 254 TTAGTCTTAGTTATTAGT 272

```

```

RESULT 6
US-09-954-531-424/c
; Sequence 424, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0

```

```
; SEQ ID NO 424
; LENGTH: 421
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)..(421)
; OTHER INFORMATION: n=a,t,g or c
US-09-954-531-424
```

```
Query Match          79.0%; Score 15.8; DB 9; Length 421;
Best Local Similarity 89.5%; Pred. No. 5.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 TTACTTCTTAGCTATTACT 19
    |||||
Db 298 TTACTTCTTAGCTATTACT 280
```

```
RESULT 7
US-09-880-107-220/c
; Sequence 220, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-MO
; CURRENT APPLICATION NUMBER: US/09/880.107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 220
; LENGTH: 421
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AAI02571
; NAME/KEY: unsure
; LOCATION: (1)..(421)
; OTHER INFORMATION: n = a or c or g or t
US-09-880-107-220
```

```
Query Match          79.0%; Score 15.8; DB 10; Length 421;
Best Local Similarity 89.5%; Pred. No. 5.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 TTACTTCTTAGCTATTACT 19
    |||||
Db 298 TTACTTCTTAGCTATTACT 280
```

```
RESULT 8
US-10-091-504-1359
; Sequence 1359, Application US/10091504
; Publication No. US20030059908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091.504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1359
; LENGTH: 492
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-091-504-1359
```

```
Query Match          79.0%; Score 15.8; DB 9; Length 492;
Best Local Similarity 89.5%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 TTACTTCTTAGCTATTACT 19
    |||||
Db 333 TTACTTCTTAGCTATTACT 351
```

```
RESULT 9
US-10-091-504-1360
; Sequence 1360, Application US/10091504
; Publication No. US20030059908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091.504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1360
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-091-504-1360
```

```
Query Match          79.0%; Score 15.8; DB 9; Length 492;
Best Local Similarity 89.5%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 TTACTTCTTAGCTATTACT 19
    |||||
Db 333 TTACTTCTTAGCTATTACT 351
```

```
RESULT 10
US-10-091-504-1361
; Sequence 1361, Application US/10091504
; Publication No. US20030059908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091.504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1361
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-091-504-1361
```

```
Query Match          79.0%; Score 15.8; DB 9; Length 492;
Best Local Similarity 89.5%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 TTACTTCTTAGCTATTACT 19
    |||||
Db 333 TTACTTCTTAGCTATTACT 351
```

```
RESULT 11
US-09-764-869-1359
; Sequence 1359, Application US/09764869
; Patent No. US20020061521A1
```

```

GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC007
CURRENT APPLICATION NUMBER: US/09/764,869
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 2442
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1359
LENGTH: 492
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-869-1359

Query Match 79.0%, Score 15.8; DB 10; Length 492;
Best Local Similarity 89.5%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAGTCTTACGTATTAGT 19
|||||
Db 333 TTAGTCTTACGTATTAGT 351

RESULT 12
US-09-764-869-1360
Sequence 1360, Application US/09764869
Patent No. US20020061521A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC007
CURRENT APPLICATION NUMBER: US/09/764,869
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 2442
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1360
LENGTH: 492
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-869-1360

Query Match 79.0%, Score 15.8; DB 10; Length 492;
Best Local Similarity 89.5%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAGTCTTACGTATTAGT 19
|||||
Db 333 TTAGTCTTACGTATTAGT 351

RESULT 13
US-09-764-869-1361
Sequence 1361, Application US/09764869
Patent No. US20020061521A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC007
CURRENT APPLICATION NUMBER: US/09/764,869
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 2442
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1361
LENGTH: 492
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-869-1361

Query Match 79.0%, Score 15.8; DB 10; Length 492;
Best Local Similarity 89.5%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Matches	17: Conservative	0: Mismatches	2: Indels	0: Gaps	0: Other
QY	1	TTAGTCTTACTATTACT	19		
Db	333	TTAGTCTTACTATTACT	351		
RESULT 14					
US-08-781-986A-994/c					
Sequence 994, Application US/08781986A					
Publication No. US20030054436A1					
GENERAL INFORMATION:					
APPLICANT: Charles Kunsch					
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences					
NUMBER OF SEQUENCES: 5255					
CORRESPONDENCE ADDRESS:					
ADDRESSEE: Human Genome Sciences, Inc.					
STREET: 9410 Key West Avenue					
CITY: Rockville					
STATE: Maryland					
COUNTRY: USA					
ZIP: 20850					
COMPUTER READABLE FORM:					
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage					
COMPUTER: HP Vectra 486/33					
OPERATING SYSTEM: MSDOS version 6.2					
SOFTWARE: ASCII Text					
CURRENT APPLICATION DATA:					
APPLICATION NUMBER: US/08/781,986A					
FILING DATE:					
CLASSIFICATION: 435					
PRIOR APPLICATION DATA:					
APPLICATION NUMBER:					
FILING DATE:					
ATTORNEY/AGENT INFORMATION:					
NAME: Benson, Bob					
REGISTRATION NUMBER: 30,446					
REFERENCE/DOCKET NUMBER: P248BP					
TELECOMMUNICATION INFORMATION:					
TELEPHONE: (301) 309-8504					
TELEFAX: (301) 309-8512					
INFORMATION FOR SEQ ID NO: 994:					
SEQUENCE CHARACTERISTICS:					
LENGTH: 551 base pairs					
TYPE: nucleic acid					
STRANDEDNESS: double					
TOPOLOGY: linear					
US-08-781-986A-994					
Query Match					
Best Local Similarity 79.0%; Score 15.8; DB 7; Length 551;					
Matches 17: Conservative 0; Mismatches 2; Indels 0; Gaps 0;					
QY	1	TTAGTCTTACTATTACT	19		
Db	476	TTAGTATTGTATTACT	458		
RESULT 15					
US-09-232-785-330					
Sequence 330, Application US/09232785					
Publication No. US20030049612A1					
GENERAL INFORMATION:					
APPLICANT: International Paper Co.					
APPLICANT: Echt, Craig S					
APPLICANT: Neilson, C. Dana					
TITLE OF INVENTION: MICROSAFELITE DNA MARKERS AND USES					
TITLE OF INVENTION: THEREOF					
FILE REFERENCE: 4481/1E180S1					
CURRENT APPLICATION NUMBER: US/09/232,785					
CURRENT FILING DATE: 1999-01-19					
PRIOR APPLICATION NUMBER: 09/232,884					
PRIOR FILING DATE: 1999-01-15					

```

; NUMBER OF SEQ ID NOS: 397
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 330
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Pinus taeda L.
; FEATURE:
; NAME/KEY: unsure
; LOCATION: n at 319.
; OTHER INFORMATION: n is a or g or c or t/u, unknown or other.
US-09-232-785-330

```

```

Query Match          79.0%; Score 15.8; DB 9; Length 573;
Best Local Similarity 89.5%; Pred. No. 5.9e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 TTAGTCTTACTTATTAGT 19
    ||||||| 1 |||||
Db 196 TTAGTCTTGTGATATTAGT 214

```

Search completed: April 6, 2003, 12:09:29  
 Job time : 60 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 10:29:22 ; Search time 168.5 seconds  
(without alignments)  
1912.600 Million cell updates/sec

Title: US-09-980-265-10

Perfect score: 20  
Sequence: 1 ttagtcttagtttagt 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 60155 seqs, 8056825618 residues

Total number of hits satisfying chosen parameters: 120310

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl\_HTG:\*  
1: gb\_hcg:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18.4	92.0	5421	1 AC019958	AC019958 Drosophila
2	18.4	92.0	8659	1 AC117270	AC117270 Drosophila
3	18.4	92.0	63673	1 AC108390	AC108390 Homo sapi
4	18.4	92.0	64195	1 AC015208	AC015208 Drosophila
5	18.4	92.0	70307	1 AC121123	AC121123 Mus muscu
6	18.4	92.0	77961	1 AC111421	AC111421 Rattus no
7	18.4	92.0	104887	1 AC104044	AC104044 Homo sapi
8	18.4	92.0	104992	1 AC005504	AC005504 Plasmodiu
9	18.4	92.0	114226	1 AC004710	AC004710 Plasmodiu
10	18.4	92.0	131347	1 AC002421	AC002421 Homo sapi
11	18.4	92.0	139205	1 AC130805	AC130805 Medicago
12	18.4	92.0	142423	1 AC109061	AC109061 Rattus no
13	18.4	92.0	153671	1 AC130755	AC130755 Rattus no
14	18.4	92.0	155877	1 AC128449	AC128449 Rattus no
15	18.4	92.0	160264	1 AC087327	AC087327 Trypanoso
16	18.4	92.0	168239	1 AC021367	AC021367 Homo sapi
17	18.4	92.0	168273	1 AC010812	AC010812 Homo sapi
18	18.4	92.0	169092	1 AC027572	AC027572 Homo sapi
19	18.4	92.0	169092	1 AC027572	AC027572 Homo sapi
20	18.4	92.0	169546	1 AC004157	AC004157 Plasmodiu
21	18.4	92.0	176235	1 AL080815	AL080815 Mus muscu
22	18.4	92.0	178875	1 AP001567	AP001567 Mus muscu
23	18.4	92.0	180035	1 AC117985	AC117985 Papio cyn
24	18.4	92.0	180685	1 AC090329	AC090329 Homo sapi
25	18.4	92.0	180783	1 AC116932	AC116932 Papio cyn
26	18.4	92.0	181655	1 AC026946	AC026946 Homo sapi
27	18.4	92.0	192337	1 AC117628	AC117628 Mus muscu
28	18.4	92.0	194339	1 AC126795	AC126795 Mus muscu
29	18.4	92.0	198345	1 AC110831	AC110831 Rattus no
30	18.4	92.0	214503	1 AL731842	AL731842 Mus muscu
31	18.4	92.0	226325	1 AL845171	AL845171 Mus muscu
32	18.4	92.0	236414	1 AC092742	AC092742 Mus muscu

33	18.4	92.0	251083	1	AC117468	Homo sapi
34	18.4	92.0	309945	1	AC122312	Mus muscu
35	17.4	87.0	937	1	AC012936	Drosophi
36	17.4	87.0	19997	1	AC015120	Drosophi
37	17.4	87.0	45817	1	AC006843	Caenorhab
38	17.4	87.0	56918	1	AC121288	Mus muscu
39	17.4	87.0	58752	1	AC008295	Drosophi
40	17.4	87.0	63037	1	AC090194	Homo sapi
41	17.4	87.0	63152	1	AC129559	Mus muscu
42	17.4	87.0	68309	1	AC123613	Mus muscu
43	17.4	87.0	73393	1	AC121234	Medicago
44	17.4	87.0	104389	1	AC009184	Drosophi
45	17.4	87.0	109740	1	AC020430	Drosophi

## ALIGNMENTS

RESULT 1	AC019958/c	AC019958	5421 bp	DNA	linear	HTG 03-JAN-2000
LOCUS	AC019958/c	Drosophila melanogaster, ***	SEQUENCING IN PROGRESS ***	in ordered		
DEFINITION	AC019958	Drosophila melanogaster, ***	SEQUENCING IN PROGRESS ***	in ordered		
PIECES	AC019958	Drosophila melanogaster, ***	SEQUENCING IN PROGRESS ***	in ordered		
ACCESSION	AC019958	GI:6664939				
VERSION	AC019958.1	GI:6664939				
KEYWORDS	HTG: HTGS_PHASE2.					
SOURCE	Drosophila melanogaster.					
ORGANISM	Drosophila melanogaster.					
REFERENCE	Adams, M. and Venter, J.C.					
AUTHORS	Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA					
JOURNAL	This sequence was identified as CDM:10211530 by the submitter. For more information on this record e-mail to fly@celera.com.					
COMMENT	* NOTE: This is a 'working draft' sequence. * This sequence will be replaced * by the finished sequence as soon as it is available and * the accession number will be preserved.					
FEATURES	Location/Qualifiers					
source	1..5421					
BASE COUNT	1558 a 1150 c 1146 g 1567 t					
ORIGIN	1558 a 1150 c 1146 g 1567 t					
Query Match	92.0% Score 18.4; DB 1; Length 5421;					
Best Local Similarity	95.0% Pred. No. 27;					
Matches	19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
Oy	1 TTAGTCTTAGTGTGTTAGTT 20					
Db	435 TTAGTCTTAGTGTGTTAGTT 416					
RESULT 2	AC117270/c	8659 bp	DNA	linear	HTG 09-APR-2002	
LOCUS	AC117270/c	Drosophila discoidium chromosome 2 map 3648613-3657270 strain				
DEFINITION	AX4.*** SEQUENCING IN PROGRESS ***	in ordered				
ACCESSION	AC117270					
VERSION	AC117270.1	GI:20087117				
KEYWORDS	HTG: HTGS_PHASE2.					
SOURCE	Drosophila discoidium.					
ORGANISM	Drosophila discoidium.					
REFERENCE	1 (bases 1 to 8659)					
AUTHORS	Gloeckner, G., Eichinger, L., Szafinski, K., Pachebat, J., Dear, P., Lehmann, R., Baumgart, C., Parra, G., April, J.F., Guigo, R., Kumpf, K., Tunggal, B., Cox, E., Quail, M.A., Platzer, M., Rosenthal, A. and					

TITLE  
JOURNAL  
REMARK  
REFERENCE  
AUTHORS  
JOURNAL  
COMMENT

Noegel, A.A.  
Sequence and Analysis of Chromosome 2 of Dictyostellium  
Unpublished  
The Dictyostellium Genome Sequencing Consortium  
2 (bases 1 to 8659)  
Baumgart, C.  
Direct Submission  
Submitted (09-APR-2002) Genome Analysis, Institute of Molecular  
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany  
CDS predictions from Genedb may contain errors. Further information  
is available from IMB Jena, Department of Genome Analysis  
(http://genome.imb-jena.de/dictyostellium/)  
and the University Cologne, Institute for Biochemistry I  
(http://www.uni-koeln.de/dictyostellium/project.shtml)  
Funding  
Agency : Deutsche Forschungsgemeinschaft (DFG).  
\* NOTE: This is a 'working draft' sequence.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

FEATURES  
source  
1. 8659  
/organism="Dictyostellium discoideum"  
/strain="AX4"  
/db\_xref="taxon:44689"  
/chromosome="2"  
/map="3648613-3657270"  
/map="3648613-3657270"  
BASE COUNT  
3236 a 1121 c 1242 g 3060 t  
ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 8659;  
Best Local Similarity 95.0%; Pred. No. 29;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTACTTCTAGTTTACTT 20  
||||| |||||||||  
Db 577 TTACTTTTACTTTTACTT 558

RESULT 3  
AC108390  
LOCUS  
DEFINITION  
AC108390 Homo sapiens chromosome 8 clone RP11-2342N23 map 8, LOW-PASS  
SEQUENCE SAMPLING.  
AC108390  
VERSION  
AC108390.1 GI:18377195  
KEYWORDS  
HTG: HTGS\_PHASE0.  
SOURCE  
Homo sapiens.  
ORGANISM  
Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 63673)  
Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
Homo sapiens chromosome 8, clone RP11-2342N23  
Unpublished  
2 (bases 1 to 63673)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N.,  
Anderson, S., Barina, N., Bastien, V., Boguslavsky, L., Bouhgalter, B.,  
Brown, A., Camarata, J., Campoliano, A., Chang, J., Chazaro, B.,  
Choquel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,  
Cooke, P., DeArrelano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,  
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,  
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
Kamat, A., Karatas, A., Kells, C., Laroque, K., Lamazares, R.,  
Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C.,  
Macdonald, P., Major, J., Margis, N., Matthews, C., McCarthy, M.,  
McEwan, P., McKenna, K., Meldrum, J., Menus, L., Mihova, T.,  
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,  
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,  
Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,  
Retta, R., Riebeck, M., Riley, R., Rise, C., Rogov, P., Roman, J.,  
Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S.,

TITLE  
JOURNAL  
COMMENT

Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
Strauss, N., Subramanian, A., Palamas, J., Testaye, S., Theodore, J.,  
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wymen, D., Ye, W.J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (27-JAN-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L13063  
Center clone name: 2342\_N23  
-----  
\* NOTE: This record contains 83 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1  
666 765: gap of 100 bp  
766 1416: contig of 651 bp in length  
1417 1516: gap of 100 bp  
1517 2160: contig of 644 bp in length  
2161 2260: gap of 100 bp  
2261 2912: contig of 652 bp in length  
2913 3012: gap of 100 bp  
3013 3681: contig of 669 bp in length  
3682 3781: gap of 100 bp  
3782 4420: contig of 639 bp in length  
4421 4520: gap of 100 bp  
4521 5167: contig of 647 bp in length  
5168 5267: gap of 100 bp  
5268 5922: contig of 655 bp in length  
5923 6022: gap of 100 bp  
6023 6676: contig of 654 bp in length  
6677 6776: gap of 100 bp  
6777 7453: contig of 677 bp in length  
7454 7553: gap of 100 bp  
7554 8255: contig of 702 bp in length  
8256 8355: gap of 100 bp  
8356 9044: contig of 689 bp in length  
9045 9144: gap of 100 bp  
9145 9805: contig of 661 bp in length  
9806 9905: gap of 100 bp  
9906 10548: contig of 643 bp in length  
10549 10648: gap of 100 bp  
10649 11310: contig of 662 bp in length  
11311 11410: gap of 100 bp  
11411 12087: contig of 677 bp in length  
12088 12187: gap of 100 bp  
12188 12872: contig of 685 bp in length  
12873 12972: gap of 100 bp  
12973 13629: contig of 657 bp in length  
13630 13729: gap of 100 bp  
13730 14406: contig of 677 bp in length  
14407 14506: gap of 100 bp  
14507 15177: contig of 671 bp in length  
15178 15277: gap of 100 bp  
15278 15930: contig of 653 bp in length  
15931 16030: gap of 100 bp



```

* 16031 16687: contig of 657 bp in length
* 16688 16787: gap of 100 bp
* 16788 17450: contig of 663 bp in length
* 17451 17550: gap of 100 bp
* 17551 18222: contig of 672 bp in length
* 18223 18322: gap of 100 bp
* 18323 18978: contig of 656 bp in length
* 18979 19078: gap of 100 bp
* 19079 19751: contig of 673 bp in length
* 19752 19851: gap of 100 bp
* 19852 20523: contig of 672 bp in length
* 20524 20623: gap of 100 bp
* 20624 21313: contig of 690 bp in length
* 21314 21413: gap of 100 bp
* 21414 22091: contig of 678 bp in length
* 22092 22191: gap of 100 bp
* 22192 22879: contig of 688 bp in length
* 22880 22979: gap of 100 bp
* 22980 23636: contig of 657 bp in length
* 23637 23736: gap of 100 bp
* 23737 24426: contig of 690 bp in length
* 24427 24526: gap of 100 bp
* 24527 25215: contig of 689 bp in length
* 25216 25315: gap of 100 bp
* 25316 25971: contig of 656 bp in length
* 25972 26071: gap of 100 bp
* 26072 26730: contig of 659 bp in length
* 26731 26830: gap of 100 bp
* 26831 27503: contig of 673 bp in length
* 27504 27603: gap of 100 bp
* 27604 28282: contig of 679 bp in length
* 28283 28382: gap of 100 bp
* 28383 29059: contig of 677 bp in length
* 29060 29159: gap of 100 bp
* 29160 29841: contig of 682 bp in length
* 29842 29941: gap of 100 bp
* 29942 30594: contig of 653 bp in length
* 30595 30694: gap of 100 bp
* 30695 31374: contig of 680 bp in length
* 31375 31474: gap of 100 bp
* 31475 32130: contig of 656 bp in length
* 32131 32230: gap of 100 bp
* 32231 32886: contig of 656 bp in length
* 32887 32986: gap of 100 bp
* 32987 33654: contig of 668 bp in length
* 33655 33754: gap of 100 bp
* 33755 34411: contig of 657 bp in length
* 34412 34511: gap of 100 bp
* 34512 35169: contig of 658 bp in length
* 35170 35269: gap of 100 bp
* 35270 35944: contig of 675 bp in length
* 35945 36044: gap of 100 bp
* 36045 36707: contig of 663 bp in length
* 36708 36807: gap of 100 bp
* 36808 37490: contig of 683 bp in length
* 37491 37590: gap of 100 bp
* 37591 38253: contig of 663 bp in length
* 38254 38353: gap of 100 bp
* 38354 39023: contig of 670 bp in length
* 39024 39123: gap of 100 bp
* 39124 39611: contig of 688 bp in length
* 39612 39911: gap of 100 bp
* 39912 40591: contig of 680 bp in length
* 40592 40691: gap of 100 bp
* 40692 41359: contig of 668 bp in length
* 41360 41459: gap of 100 bp
* 41460 42122: contig of 663 bp in length
* 42123 42222: gap of 100 bp
* 42223 42898: contig of 676 bp in length
* 42899 42998: gap of 100 bp
* 42999 43685: contig of 687 bp in length
* 43686 43785: gap of 100 bp
* 43786 44450: contig of 665 bp in length

```

```

Query Match          92.0%: Score 18.4; DB 1; Length 63673;
Best Local Similarity 95.0%: Pred. No. 36;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 TTAGTCTTACGTTTACTT 20
DB 44734 TTAGTTTTCAGTTTACTT 44753

```

```

RESULT 4
AC015208/c 64195 bp DNA linear HTG 16-NOV-1999
DEFINITION Drosophila melanogaster. *** SEQUENCING IN PROGRESS ***; in ordered
pieces.
ACCESSION AC015208
VERSION AC015208.1 GI:6436127
KEYWORDS HTG; HTGS_PHASE2.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 64195)
REFERENCE Adams, M. and Venter, J.C.
AUTHORS Direct Submission
TITLE Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,
JOURNAL Rockville, MD, USA
COMMENT This sequence was identified as CDM:10211252 by the submitter.
For further information on this sequence e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

```

```

FEATURES
source 1..64195
location/Qualifiers

```

```

BASE COUNT 17633 a 14801 c 14632 g 17129 t
ORIGIN

```

```

Query Match          92.0%: Score 18.4; DB 1; Length 64195;
Best Local Similarity 95.0%: Pred. No. 36;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 TTAGTCTTACGTTTACTT 20
DB 4393 TTAGTTCATGTTTACTT 4374

```

```

RESULT 5

```

AC121123 70307 bp DNA linear HTG 15-MAY-2002  
LOCUS Mus musculus clone RP23-474E16, LOW-PASS SEQUENCE SAMPLING.  
DEFINITION AC121123  
ACCESSION AC121123  
VERSION AC121123.1 GI:20800260  
KEYWORDS HTG; HTGS\_PHASE0.  
SOURCE Mus musculus.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 70307)  
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
JOURNAL Mus musculus, clone RP23-474E16  
TITLE Unpublished  
REFERENCE 2 (bases 1 to 70307)  
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,  
Anderson, S., Bastien, V., Bloom, T., Boguslavsky, L.,  
Boukhalter, B., Brown, A., Camarata, J., Campolano, A., Chang, J.,  
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,  
Cook, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J., Dodge, S.,  
Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,  
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-pierre, N.,  
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamazares, R.,  
Landers, T., Lehoczký, J., Levine, R., Lindblad-Toh, K., Liu, G.,  
Maclean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C.,  
McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menues, L.,  
Mhova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,  
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,  
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,  
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
Strauss, N., Subramanian, A., Talamas, J., Testafaye, S., Theodore, J.,  
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (15-MAY-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
COMMENT All repeats were identified using RepeatMasker:  
Smith, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: MIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [submissions@genome.wi.mit.edu](mailto:submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L22814  
Center clone name: 474\_E\_16  
-----  
\* NOTE: This record contains 87 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.  
\* 1 702: contig of 702 bp in length  
\* 703 802: gap of 100 bp  
\* 803 1523: contig of 721 bp in length  
\* 1524 1623: gap of 100 bp  
\* 1624 2341: contig of 718 bp in length  
\* 2342 2441: gap of 100 bp  
\* 2442 3142: contig of 701 bp in length  
\* 3143 3242: gap of 100 bp  
\* 3243 3945: contig of 703 bp in length  
\* 3946 4045: gap of 100 bp  
-----  
4046 4737: contig of 692 bp in length  
\* 4738 4837: gap of 100 bp  
\* 4838 5501: contig of 664 bp in length  
\* 5502 5601: gap of 100 bp  
\* 5602 6315: contig of 714 bp in length  
\* 6316 6415: gap of 100 bp  
\* 6416 7138: contig of 723 bp in length  
\* 7139 7238: gap of 100 bp  
\* 7239 7964: contig of 726 bp in length  
\* 7965 8064: gap of 100 bp  
\* 8065 8765: contig of 701 bp in length  
\* 8766 8865: gap of 100 bp  
\* 8866 9578: contig of 713 bp in length  
\* 9579 9678: gap of 100 bp  
\* 9679 10404: contig of 726 bp in length  
\* 10405 10504: gap of 100 bp  
\* 10505 11206: contig of 702 bp in length  
\* 11207 11306: gap of 100 bp  
\* 11307 12002: contig of 696 bp in length  
\* 12003 12102: gap of 100 bp  
\* 12103 12806: contig of 704 bp in length  
\* 12807 12906: gap of 100 bp  
\* 12907 13614: contig of 708 bp in length  
\* 13615 13714: gap of 100 bp  
\* 13715 14426: contig of 712 bp in length  
\* 14427 14526: gap of 100 bp  
\* 14527 15253: contig of 727 bp in length  
\* 15254 15353: gap of 100 bp  
\* 15354 16083: contig of 730 bp in length  
\* 16084 16183: gap of 100 bp  
\* 16184 16877: contig of 694 bp in length  
\* 16878 16977: gap of 100 bp  
\* 16979 17694: contig of 717 bp in length  
\* 17695 17794: gap of 100 bp  
\* 17795 18510: contig of 716 bp in length  
\* 18511 18610: gap of 100 bp  
\* 18611 19309: contig of 699 bp in length  
\* 19310 19409: gap of 100 bp  
\* 19410 20118: contig of 709 bp in length  
\* 20119 20218: gap of 100 bp  
\* 20219 20909: contig of 691 bp in length  
\* 20910 21009: gap of 100 bp  
\* 21010 21702: contig of 693 bp in length  
\* 21703 21802: gap of 100 bp  
\* 21803 22526: contig of 724 bp in length  
\* 22527 22626: gap of 100 bp  
\* 22627 23333: contig of 707 bp in length  
\* 23334 23433: gap of 100 bp  
\* 23434 24154: contig of 721 bp in length  
\* 24155 24254: gap of 100 bp  
\* 24255 24960: contig of 726 bp in length  
\* 24961 25080: gap of 100 bp  
\* 25081 25785: contig of 705 bp in length  
\* 25786 25885: gap of 100 bp  
\* 25886 26609: contig of 724 bp in length  
\* 26610 26709: gap of 100 bp  
\* 26710 27439: contig of 730 bp in length  
\* 27440 27539: gap of 100 bp  
\* 27540 28243: contig of 704 bp in length  
\* 28244 28343: gap of 100 bp  
\* 28344 29032: contig of 689 bp in length  
\* 29033 29132: gap of 100 bp  
\* 29133 29816: contig of 684 bp in length  
\* 29817 29916: gap of 100 bp  
\* 29917 30611: contig of 695 bp in length  
\* 30612 30711: gap of 100 bp  
\* 30712 31407: contig of 696 bp in length  
\* 31408 31507: gap of 100 bp  
\* 31508 32234: contig of 727 bp in length  
\* 32235 32334: gap of 100 bp  
\* 32335 33066: contig of 732 bp in length  
\* 33067 33166: gap of 100 bp  
\* 33167 33867: contig of 701 bp in length

```

* 3368 33967: gap of 100 bp
* 33968 34682: contig of 715 bp in length
* 34683 34782: gap of 100 bp
* 34783 35508: contig of 726 bp in length
* 35509 35608: gap of 100 bp
* 35609 36286: contig of 678 bp in length
* 36287 36386: gap of 100 bp
* 36387 37094: contig of 708 bp in length
* 37095 37194: gap of 100 bp
* 37195 37885: contig of 691 bp in length
* 37886 37985: gap of 100 bp
* 37986 38687: contig of 702 bp in length
* 38688 38787: gap of 100 bp
* 38788 39487: contig of 700 bp in length
* 39488 39587: gap of 100 bp
* 39588 40290: contig of 703 bp in length
* 40291 40390: gap of 100 bp
* 40391 41120: contig of 730 bp in length
* 41121 41220: gap of 100 bp
* 41221 41948: contig of 728 bp in length
* 41949 42048: gap of 100 bp
* 42049 42756: contig of 708 bp in length
* 42757 42856: gap of 100 bp
* 42857 43581: contig of 725 bp in length
* 43582 43681: gap of 100 bp
* 43682 44406: contig of 725 bp in length
* 44407 44506: gap of 100 bp
* 44507 45207: contig of 701 bp in length
* 45208 45307: gap of 100 bp
* 45308 46004: contig of 697 bp in length
* 46005 46104: gap of 100 bp
* 46105 46787: contig of 683 bp in length
* 46788 46887: gap of 100 bp
* 46888 47572: contig of 685 bp in length
* 47573 47672: gap of 100 bp
* 47673 48391: contig of 719 bp in length
* 48392 48491: gap of 100 bp
* 48492 49198: contig of 707 bp in length
* 49199 49298: gap of 100 bp
* 49299 50024: contig of 726 bp in length
* 50025 50124: gap of 100 bp
* 50125 50857: contig of 733 bp in length
* 50858 50957: gap of 100 bp
* 50958 51684: contig of 727 bp in length
* 51685 51784: gap of 100 bp
* 51785 52487: contig of 703 bp in length
* 52488 52587: gap of 100 bp
* 52588 53313: contig of 726 bp in length
* 53314 53413: gap of 100 bp
* 53414 54132: contig of 719 bp in length
* 54133 54232: gap of 100 bp
* 54233 54949: contig of 717 bp in length
* 54950 55049: gap of 100 bp
* 55050 55726: contig of 677 bp in length
* 55727 55826: gap of 100 bp

```

Query Match 92.0%: Score 18.4; DB 1; Length 70307;  
 Best Local Similarity 95.0%: Pred. No. 36;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 TTAGTCTTCTAGTTTCTAGTT 20  
 Db 20430 TCAGTCTTCTAGTTTCTAGTT 20449

RESULT 6  
 AC111421/c 77961 bp DNA linear HTG 13-JUL-2002  
 LOCUS Rattus norvegicus clone CH230-45G18, \*\*\* SEQUENCING IN PROGRESS  
 DEFINITION \*\*\* 44 unordered pieces.  
 ACCESSION AC111421  
 AC111421.2 GI:21735962  
 KEYWORDS HTG; HTGS\_PHASE1.

SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE 1 (bases 1 to 77961)  
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
 Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Ayele,M., Banks,T.,  
 Barbara,J., Benton,J., Bimage,K., Blankenburg,K., Bonini,D.,  
 Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,  
 Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
 Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,  
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
 Davila,M.L., Davis,C., Day-Carroll,L., Dedrich,D.A.,  
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durkin,K.J.,  
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
 Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,  
 Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,  
 Homsl,F., Howard,S., Huber,J., Huliyk,S., Hume,J., Jackson,L.E.,  
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
 Karlsson,E., Kelly,S., Khan,U., Kling,L., Korvah,J., Kovar,C.,  
 Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
 Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louisedge,H.,  
 Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
 Massey,E., Mawhiny,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M.,  
 Miner,G., Miner,Z., Mitchell,T., Mohabadi,K., Morgan,M., Morris,S.,  
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
 Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okunogun,G.,  
 Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,D., Perez,L.,  
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
 Rives,M., Rojas,A., Rojudoan,I., Rolfe,M., Ruiz,S., Savery,G.,  
 Scherer,S., Scott,G., Shen,H., Shoshbari,N., Sisson,I.,  
 Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,  
 Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
 Tamey,J., Taylor,C., Taylor,T., Telitod,B., Thomas,N., Thomas,S.,  
 Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,  
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,  
 Williams,G., Williamson,A., Wleczky,R., Woodson,S., Worley,K.,  
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 Weinstein,G. and Gibbs,R.

TITLE Direct Submission  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 77961)  
 AUTHOR S. Worley, K.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-FEB-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 77961)  
 REFERENCE Direct Submission  
 AUTHOR S. Worley, K.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Jul 12, 2002 this sequence version replaced gi:18701185.  
 COMMENT ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: http://www.hgsc.bcm.tmc.edu/  
 Contact: hgsc-help@bcm.tmc.edu  
 Project information  
 Center project name: GMG  
 Center Clone name: CH230-45G18  
 Summary Statistics  
 Sequencing vector: Plasmid  
 Chemistry: Dye-terminator Big Dye: 100% of reads  
 Assembly program: Phrap: version 0.990329



McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J.,  
Meneu, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,  
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,  
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,  
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
Topham, K., Travers, M., Travis, N., Triggiani, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
Submitted (01-DEC-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 104887)

Barron, B., Nusbbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,  
Barra, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B.,  
Camata, J., Chang, J., Chazaro, B., Choepel, Y., Collamore, A.,  
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,  
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,  
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,  
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A.,  
Karats, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,  
Lin, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,  
McCarthy, M., Meldrum, J., Meneu, L., Mihova, T., Mlenga, V.,  
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,  
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,  
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,  
Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P.,  
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,  
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,  
Zembek, L., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
COMMENT  
Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Aug 6, 2002 this sequence version replaced g1:18643594.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence.submissions@genome.wi.mit.edu  
Project Information  
Center project name: 2219\_A\_7  
Center clone name: 2219\_A\_7

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 2 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 1 27562: contig of 27562 bp in length  
\* 27563 27662: gap of 100 bp  
\* 27663 104887: contig of 77225 bp in length.  
Location/Qualifiers

FEATURES  
SOURCE

1..104887  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8"

BASE COUNT  
ORIGIN  
31026 a 20095 c 20706 g 32141 t 915 others

Query Match

92.0%; Score 18.4; DB 1; length 104887;

Best Local Similarity 95.0%; Pred. No. 38;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTACGTTTACGTT 20  
Db 35479 TTAGTCTTACGTTTACGTT 35498

RESULT 8

AC005504/C

LOCUS

DEFINITION

AC005504

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (21-AUG-1998) Stanford DNA Sequencing and Technology  
Center, Stanford University, 855 California Avenue, Palo Alto, CA  
94304, USA

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 3 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 58642: contig of 58642 bp in length  
\* 58643 58843: gap of unknown length  
\* 58843 91011: contig of 32169 bp in length  
\* 91012 91211: gap of unknown length  
\* 91212 104992: contig of 13781 bp in length.  
Location/Qualifiers

1..104992  
/organism="Plasmodium falciparum"  
/db\_xref="taxon:5833"  
/chromosome="12"

BASE COUNT  
ORIGIN  
44286 a 9326 c 9564 g 41411 t 405 others

Query Match  
Best Local Similarity 95.0%; Pred. No. 38;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTACGTTTACGTT 20  
Db 57851 TTAGTCTTACGTTTACGTT 57832

RESULT 9

AC004710/C

LOCUS

DEFINITION

AC004710

VERSION

KEYWORDS

SOURCE

ORGANISM

114226 bp DNA linear HTG 01-APR-1999  
Plasmodium falciparum chromosome 12, \*\*\* SEQUENCING IN PROGRESS  
\*\*\*, 2 unordered pieces.

AC004710.3 GI:4558577  
HTG: HTGS\_PHASE1  
Plasmodium falciparum  
Plasmodium falciparum  
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
1 (bases 1 to 114226)

AUTHORS Hyman, R.W., Fung, E.L., Qin, F., Tamaki, T., Kurd, O.B., Conway, A.B. and Davis, R.W.

TITLE Plasmodium falciparum 3D7 chromosome 12

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 114226)

AUTHORS Hyman, R.W., Qin, F., Fung, E.L., Conway, A.B. and Davis, R.W.

TITLE Direct Submission

JOURNAL Submitted (21-MAY-1998) Stanford DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA

COMMENT On Apr 2, 1999 this sequence version replaced gi:4337166. \* NOTE: This is a 'working draft' sequence. It currently \* consists of 2 contigs. The true order of the pieces \* is not known and their order in this sequence record is \* arbitrary. Gaps between the contigs are represented as \* runs of N, but the exact sizes of the gaps are unknown. \* This record will be updated with the finished sequence \* as soon as it is available and the accession number will \* be preserved.

FEATURES

source 1. 114226

location/Qualifiers

1. 114226

/organism="Plasmodium falciparum"

/db\_xref="taxon:5833"

/chromosome="12"

BASE COUNT 44338 a 11954 c 10111 g 47622 t 201 others

ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 114226;

Best Local Similarity 95.0%; Pred. No. 38;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTACTTCTAGTTTGGTT 20

Db 92278 TTACTTCTAGTTTGGTT 92259

RESULT 10

AC002421 131347 bp DNA linear HTG 13-JUN-2002

LOCUS Homo sapiens chromosome X clone PMXD1, \*\*\* SEQUENCING IN PROGRESS

DEFINITION \*\*\* 4 unordered pieces.

AC002421

VERSION AC002421.2 GI:21405641

KEYWORDS HTG; HTGS\_PHASE1.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 131347)

AUTHORS Chen, E., Brownstein, B.H., States, D.J., Schlessinger, D. and Mazzarella, R.

TITLE Direct Submission

JOURNAL Unpublished (1997)

REFERENCE 2 (bases 1 to 131347)

AUTHORS Brownstein, B.H., States, D.J. and Mazzarella, R.

TITLE Direct Submission

JOURNAL Submitted (12-AUG-1997) Center for Genetics in Medicine, Box 8232, Washington University School of Medicine, 4566 Scott Avenue, St. Louis, MO 63110, USA

COMMENT On Jun 13, 2002 this sequence version replaced gi:2323248. \* Current status of this project is available at: \* [http://genome.wustl.edu/cgm/seq\\_projects.html](http://genome.wustl.edu/cgm/seq_projects.html) \* Submitted by: \* Elison Chen, \* Advanced Center for Genetic Technology, \* Applied Biosystems Division of Perlin Elmer Corp., \* 850 Lincoln Center Drive, \* Foster City, CA 94404 USA \* e-mail: [elison@genseq.apliddio.com](mailto:elison@genseq.apliddio.com)

and

Buddy Brownstein, \* Center for Genetics in Medicine, \* Washington University School of Medicine, Box 8232 \* 4566 Scott Avenue \* St. Louis, MO 63110, USA \* e-mail: [buddy@genetics.wustl.edu](mailto:buddy@genetics.wustl.edu)

and

David J. States, \* Institute for Biomedical Computing \* Washington University in St. Louis \* 700 South Euclid Ave. \* St. Louis, MO 63108 USA \* e-mail: [states@bioc.wustl.edu](mailto:states@bioc.wustl.edu).

\* NOTE: This is a 'working draft' sequence. It currently \* consists of 4 contigs. The true order of the pieces \* is not known and their order in this sequence record is \* arbitrary. Gaps between the contigs are represented as \* runs of N, but the exact sizes of the gaps are unknown. \* This record will be updated with the finished sequence \* as soon as it is available and the accession number will \* be preserved.

FEATURES

source 1. 131347

location/Qualifiers

1. 131347

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="X"

/clone="PMXD1"

BASE COUNT 40590 a 23040 c 23923 g 43494 t 300 others

ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 131347;

Best Local Similarity 95.0%; Pred. No. 39;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTACTTCTAGTTTGGTT 20

Db 45394 TTACTTCTAGTTTGGTT 45413

RESULT 11

AC130805 139205 bp DNA linear HTG 21-AUG-2002

LOCUS Medicago truncatula clone mth2-9n1, WORKING DRAFT SEQUENCE, 7

DEFINITION unordered pieces.

AC130805

VERSION AC130805.2 GI:22325011

KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.

SOURCE barrel medic.

ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

REFERENCE 1 (bases 1 to 139205)

AUTHORS Shaul, S., Lin, S., Dixon, R., May, G., Summer, L., Gonzales, B., Cook, D., Kim, D. and Roe, B.A.

TITLE Medicago truncatula BAC Clone mth2-9n1

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 139205)

AUTHORS Shaul, S., Lin, S., Dixon, R., May, G., Summer, L., Gonzales, B., Cook, D., Kim, D. and Roe, B.A.

TITLE Direct Submission

JOURNAL Submitted (14-AUG-2002) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
REFERENCE 3 (bases 1 to 139205)  
AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,  
Cook,D., Kim,D. and Roe,B.A.  
TITLE Direct Submission  
JOURNAL Submitted (21-AUG-2002) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
COMMENT On Aug 20, 2002 this sequence version replaced gi:22218497.  
----- Genome Center  
Center: Department Of Chemistry And Biochemistry  
The University Of Oklahoma  
Center code:UOKNOR  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 7 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 2045: contig of 2045 bp in length  
\* 2046 2145: gap of unknown length  
\* 2146 5410: contig of 3265 bp in length  
\* 5411 5510: gap of unknown length  
\* 5511 11789: contig of 6279 bp in length  
\* 11790 11889: gap of unknown length  
\* 11890 18821: contig of 6932 bp in length  
\* 18822 18921: gap of unknown length  
\* 18922 35712: contig of 16791 bp in length  
\* 35713 35812: gap of unknown length  
\* 35813 65594: contig of 29782 bp in length  
\* 65595 139205: contig of 73511 bp in length.  
\* 65695  
\* Location/Qualifiers  
\* 1. 139205  
\* /organism="Medicago truncatula"  
\* /db\_xref="taxon:3880"  
\* /clone="mth2-9n1"  
\* /clone\_lib="Medicago truncatula BAC library H2"  
\* /clone\_id="23034 c 21787 g 46196 t 601 others  
BASE COUNT 47587 a 23034 c 21787 g 46196 t 601 others  
ORIGIN  
Query Match 92.0%; Score 18.4; DB 1; Length 139205;  
Best Local Similarity 95.0%; Pred. No. 39;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TTAGTCTTACGTTTACTT 20  
|||||  
DB 96871 TTAGTCTTACGTTTACTT 96852  
RESULT 12  
AC109061 142423 bp DNA linear HTG 17-JUL-2002  
LOCUS Rattus norvegicus clone CH230-101L7, \*\*\* SEQUENCING IN PROGRESS  
DEFINITION \*\*\* 69 unordered pieces.  
ACCESSION AC109061  
VERSION AC109061.3 GI:21744138  
KEYWORDS HTG: HTGS, PHASEL.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 142423)  
AUTHORS Munry,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,  
Barbata,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,  
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,

Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Daya,M.L., Davis,C., Davy-Carroll,J., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Dublin,K.J.,  
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flag,N., Ford,J., Foster,P., Frantz,P.,  
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hogues,M., Hollway,C., Hollins,B.,  
Honsi,F., Howard,S., Huber,J., Huliy,K., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlsom,E., Kelly,S., Khan,U., King,L., Kovach,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,  
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhney,E., Mcleod,M.P., Meador,M., Mel,G., Metker,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Mosier,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nkokenkwo,S., Ogih,M., Okwuonu,G.,  
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pu,L., Quiles,M., Ren,Y.,  
Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,  
Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I.,  
Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Swalek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansley,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Usmani,K., Vasquez,L., Vera,Y., Villalon,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,  
Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G. and Gibbs,R.  
Direct Submission  
Unpublished  
2 (bases 1 to 142423)  
Worley,K.C.  
Direct Submission  
Submitted (03-FEB-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 142423)  
Worley,K.C.  
Direct Submission  
Submitted (17-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 13, 2002 this sequence version replaced gi:18846789.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GPV  
Center clone name: CH230-101L7  
----- Summary Statistics  
Sequencing vector: Plasmid  
Chemistry: Dye-terminator Big Dye: 100% of reads  
Assembly program: Phrap: version 0.990329  
Consensus quality: 84036 bases at least Q40  
Consensus quality: 88564 bases at least Q30  
Consensus quality: 92346 bases at least Q20  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_drift\_data.html).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 69 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 1004: contig of 1004 bp in length  
\* 1005 1104: gap of unknown length  
\* 1105 2142: contig of 1038 bp in length  
\* 2143 2242: gap of unknown length  
\* 2243 3287: contig of 1045 bp in length  
\* 3288 3388: gap of unknown length  
\* 3388 4626: contig of 1239 bp in length  
\* 4627 4727: gap of unknown length  
\* 4727 5837: contig of 1111 bp in length  
\* 5838 5937: gap of unknown length  
\* 5938 7447: contig of 1510 bp in length  
\* 7448 7547: gap of unknown length  
\* 7548 8871: contig of 1324 bp in length  
\* 8872 8971: gap of unknown length  
\* 8972 10001: contig of 1030 bp in length  
\* 10002 10101: gap of unknown length  
\* 10102 11384: contig of 1283 bp in length  
\* 11385 11484: gap of unknown length  
\* 11485 12538: contig of 1054 bp in length  
\* 12539 12638: gap of unknown length  
\* 12639 13875: contig of 1237 bp in length  
\* 13876 13975: gap of unknown length  
\* 13976 15308: contig of 1533 bp in length  
\* 15309 15608: gap of unknown length  
\* 15609 16953: contig of 1345 bp in length  
\* 16954 17053: gap of unknown length  
\* 17054 18233: contig of 1180 bp in length  
\* 18234 18333: gap of unknown length  
\* 18334 19344: contig of 1011 bp in length  
\* 19345 19444: gap of unknown length  
\* 19445 20666: contig of 1222 bp in length  
\* 20667 20766: gap of unknown length  
\* 20767 22253: contig of 1487 bp in length  
\* 22254 22353: gap of unknown length  
\* 22354 23999: contig of 1646 bp in length  
\* 24000 24099: gap of unknown length  
\* 24100 25396: contig of 1297 bp in length  
\* 25397 25496: gap of unknown length  
\* 25497 27117: contig of 1621 bp in length  
\* 27118 27217: gap of unknown length  
\* 27218 28890: contig of 1673 bp in length  
\* 28891 28990: gap of unknown length  
\* 28991 31178: contig of 2188 bp in length  
\* 31179 31278: gap of unknown length  
\* 31279 32681: contig of 1403 bp in length  
\* 32682 32781: gap of unknown length  
\* 32782 34782: contig of 2001 bp in length  
\* 34783 34882: gap of unknown length  
\* 34883 35991: contig of 1109 bp in length  
\* 35992 36091: gap of unknown length  
\* 36092 37963: contig of 1872 bp in length  
\* 37964 38063: gap of unknown length  
\* 38064 40364: contig of 2301 bp in length  
\* 40365 40464: gap of unknown length  
\* 40465 41813: contig of 1349 bp in length  
\* 41814 41913: gap of unknown length  
\* 41914 43241: contig of 1328 bp in length  
\* 43242 43341: gap of unknown length  
\* 43342 44828: contig of 1487 bp in length  
\* 44829 44928: gap of unknown length  
\* 44929 46483: contig of 1555 bp in length  
\* 46484 46583: gap of unknown length  
\* 46584 47891: contig of 1308 bp in length  
\* 47892 47991: gap of unknown length  
\* 47992 49446: contig of 1455 bp in length  
\* 49447 49546: gap of unknown length  
\* 49547 51449: contig of 1903 bp in length  
\* 51450 51549: gap of unknown length  
\* 51550 53745: contig of 2196 bp in length

\* 53746 53845: gap of unknown length  
\* 53846 56029: contig of 2184 bp in length  
\* 56030 56129: gap of unknown length  
\* 56130 57995: contig of 1866 bp in length  
\* 57996 58095: gap of unknown length  
\* 58096 59786: contig of 1691 bp in length  
\* 59787 59886: gap of unknown length  
\* 59887 61264: contig of 1378 bp in length  
\* 61265 61364: gap of unknown length  
\* 61365 62623: contig of 1259 bp in length  
\* 62624 62723: gap of unknown length  
\* 62724 64854: contig of 2131 bp in length  
\* 64855 64954: gap of unknown length  
\* 64955 67347: contig of 2393 bp in length  
\* 67348 67348: gap of unknown length  
\* 67349 69033: contig of 1586 bp in length  
\* 69034 69133: gap of unknown length  
\* 69134 70383: contig of 1250 bp in length  
\* 70384 70484: gap of unknown length  
\* 70485 72010: contig of 1527 bp in length  
\* 72011 72110: gap of unknown length  
\* 72111 74481: contig of 2371 bp in length  
\* 74482 74581: gap of unknown length  
\* 74582 76837: contig of 2256 bp in length  
\* 76838 76937: gap of unknown length  
\* 76938 79589: contig of 2652 bp in length  
\* 79590 79689: gap of unknown length  
\* 79690 82043: contig of 2354 bp in length  
\* 82044 82143: gap of unknown length  
\* 82144 84899: contig of 2756 bp in length  
\* 84900 84999: gap of unknown length  
\* 85000 87010: contig of 2011 bp in length  
\* 87011 87110: gap of unknown length  
\* 87111 89409: contig of 2299 bp in length  
\* 89410 89509: gap of unknown length  
\* 89510 91888: contig of 2379 bp in length  
\* 91889 91989: gap of unknown length  
\* 91990 94785: contig of 2797 bp in length

Query Match 92.0%; Score 18.4; DB 1; Length 142423;  
Best Local Similarity 95.0%; Pred. No. 39;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTAGTTCTTAGTTTACTT 20  
Db 86537 TTAGTTCTTAGTTTACTT 86556

## RESULT 13

AC130755

LOCUS

DEFINITION

Rattus norvegicus clone CH230-10D6, \*\*\* SEQUENCING IN PROGRESS \*\*\*  
77 unordered pieces.

AC130755

AC130755.1 GI:22218406

VERSION

HTG: HTGS\_PHASE1.

KEYWORDS

SOURCE

ORGANISM

Norway rat.

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 153671)

REFERENCE

AUTHORS

Muzny,D.,Marle., Metzker,M.,Lee., Abramzon,S., Adams,C., Alder,J.,  
Allen,C., Allen,H., Alspbrooks,S., Amlin,A., Anguiano,D.,  
Anylebech,V., Ayagi,A., Ayodeji,M., Baca,E., Baden,H.,  
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,  
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,  
Bryant,N., Bunay,C., Burch,P., Butrell,K., Calderon,E.,  
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,  
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chu,J.,  
Cleveland,C., Cockrell,R., Cox,C., Coyler,M., Cree,A., D'Souza,L.,  
Davila,M.L., Davis,C., Day-Carroll,L., De Anda,C., Dederich,D.,  
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinn,H., Divya,K.,



Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flanagan, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gerogeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hughes, M., Hollins, B., Howells, S., Hulys, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpaty, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowals, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensushewa, L., Louisedge, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawlinney, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackelme, O., Okunolu, G., Olarunpusagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, M., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X., Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tilney, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willison, R., Wleciyk, R., Woden, H., Morley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, D., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G. and Gibbs, R.A.

Unpublished  
Direct Submission  
2 (bases 1 to 153671)  
Rat Genome Sequencing Consortium.  
Submitted (14-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

----- Genome Center -----  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information -----  
Project name: GJYZ  
Center project name: CH230-10D6  
----- Summary Statistics -----  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 72132 bases at least Q40  
Consensus quality: 78027 bases at least Q30  
Consensus quality: 81644 bases at least Q20

----- NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank/draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html)).  
NOTE: This is a 'working draft' sequence. It currently  
consists of 77 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1  
1089  
1188: contig of 1088 bp in length  
1188: gap of unknown length

1189  
2277: contig of 1089 bp in length  
2378  
2378: gap of unknown length  
3278  
3469: contig of 1092 bp in length  
3470  
3569: gap of unknown length  
3570  
4600: contig of 1031 bp in length  
4601  
4700: gap of unknown length  
4701  
5725: contig of 1025 bp in length  
5726  
5825: gap of unknown length  
5826  
6894: contig of 1069 bp in length  
6895  
6994: gap of unknown length  
6995  
8368: contig of 1374 bp in length  
8369  
8468: gap of unknown length  
8469  
9821: contig of 1353 bp in length  
9822  
9922: gap of unknown length  
9922  
11138: contig of 1217 bp in length  
11139  
11238: gap of unknown length  
11239  
12775: contig of 1537 bp in length  
12776  
12875: gap of unknown length  
12876  
13883: contig of 1008 bp in length  
13884  
13983: gap of unknown length  
13984  
15201: contig of 1218 bp in length  
15202  
15301: gap of unknown length  
15302  
16748: contig of 1447 bp in length  
16749  
16848: gap of unknown length  
16849  
18207: contig of 1359 bp in length  
18208  
18307: gap of unknown length  
18308  
20081: contig of 1774 bp in length  
20082  
20181: gap of unknown length  
20182  
21605: contig of 1424 bp in length  
21606  
21705: gap of unknown length  
21706  
23088: contig of 1333 bp in length  
23089  
23188: gap of unknown length  
23189  
24314: contig of 1126 bp in length  
24315  
24414: gap of unknown length  
24415  
25833: contig of 1419 bp in length  
25834  
25933: gap of unknown length  
25934  
27438: contig of 1505 bp in length  
27439  
27538: gap of unknown length  
27539  
28737: contig of 1199 bp in length  
28738  
28837: gap of unknown length  
28838  
30058: gap of unknown length  
30059  
30158: gap of unknown length  
30159  
31312: contig of 1154 bp in length  
31313  
31412: gap of unknown length  
31413  
32876: contig of 1464 bp in length  
32877  
32976: gap of unknown length  
32977  
34532: contig of 1556 bp in length  
34533  
34632: gap of unknown length  
34633  
35959: contig of 1327 bp in length  
35960  
37062: gap of unknown length  
37062  
37162: contig of 1003 bp in length  
37163  
37287: gap of unknown length  
37287  
38287: contig of 1125 bp in length  
38288  
38387: gap of unknown length  
38388  
39705: contig of 1318 bp in length  
39706  
39805: gap of unknown length  
39805  
40990: gap of unknown length  
40990  
41090: contig of 1185 bp in length  
41091  
42586: gap of unknown length  
42586  
42686: contig of 1456 bp in length  
42687  
42686: gap of unknown length  
42687  
43798: contig of 1112 bp in length  
43799  
43898: gap of unknown length  
43899  
45415: contig of 1517 bp in length  
45416  
45515: gap of unknown length  
45515  
46835: gap of unknown length  
46835  
46935: contig of 1320 bp in length  
46935  
48361: gap of unknown length  
48361  
48461: contig of 1426 bp in length  
48462  
48461: gap of unknown length  
48462  
50826: contig of 2365 bp in length  
50827  
50926: gap of unknown length  
50927  
52239: contig of 1313 bp in length  
52239  
52340  
52240: gap of unknown length  
52240  
53729: contig of 1390 bp in length  
53729

```

* 53730 53829: gap of unknown length
* 53830 55495: contig of 1666 bp in length
* 55496 55595: gap of unknown length
* 55596 57455: contig of 1860 bp in length
* 57456 57555: gap of unknown length
* 57556 59157: contig of 1602 bp in length
* 59158 59258: gap of unknown length
* 59259 60521: contig of 1264 bp in length
* 60522 60621: gap of unknown length
* 60622 61793: contig of 1172 bp in length
* 61794 61893: gap of unknown length
* 61894 63941: contig of 2048 bp in length
* 63942 64041: gap of unknown length
* 64042 66372: contig of 2331 bp in length
* 66373 66472: gap of unknown length
* 66473 67931: contig of 1459 bp in length
* 67932 68031: gap of unknown length
* 68032 70210: contig of 2179 bp in length
* 70211 70310: gap of unknown length
* 70311 72062: contig of 1752 bp in length
* 72063 72162: gap of unknown length
* 72163 73317: contig of 1155 bp in length
* 73318 74417: gap of unknown length
* 74418 74757: contig of 1340 bp in length
* 74758 74857: gap of unknown length
* 74858 76526: contig of 1669 bp in length
* 76527 76626: gap of unknown length
* 76627 77979: contig of 1353 bp in length
* 77980 78079: gap of unknown length
* 78080 80739: contig of 2660 bp in length
* 80740 80839: gap of unknown length

```

```

Query Match
Best Local Similarity 92.0%; Score 18.4; DB 1; Length 153671;
Pred. No. 39;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 TTACTTCTACTTTTACTT 20
Db 108998 TTACTTCTACTTTTACTT 109017

```

```

RESULT 14
AC128449
LOCUS
DEFINITION
Rattus norvegicus clone CH230-117D11, *** SEQUENCING IN PROGRESS
***, 85 unordered pieces.

```

```

ACCESSION AC128449.1 GI:21909143
VERSION AC128449.1
KEYWORDS HTGS_PHASE1.
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus.

```

```

REFERENCE
AUTHORS

```

```

1 (bases 1 to 155877)
Muzny,D.M., Adams,C., Adio-Oduola,B., Alt-osman,F.R., Allen,C.,
Alshrooqs,S.L., Amaralunga,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbora,J., Benton,J., Bimge,K., Blankenburg,K., Bonnin,D.,
Bouch,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carton,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dahorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dihn,H.H.,
Doutwalte,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferriguto,D., Flagg,N., Ford,J., Foster,P., Franz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homsli,F., Howard,S., Huber,D., Hulyk,S., Hume,J., Jackson,L.E.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

```

Jacobson,B., Jia,Y., Johnson,R., Jolyet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.,
Li,J., Li,Z., Licharge,O., Lien,C., Liu,J., Liu,W., Louised,H.,
Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Meshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenwo,S., Ogulu,M., Okunolu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojokan,I., Rolfe,M., Ruiz,S., Saverly,G.,
Scherer,S., Scott,G., Shen,H., Shooshari,N., Sisson,I.,
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G., and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 155877)
Worley,K.C.
Direct Submission
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GYWL
Center clone name: CH230-117D11
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap: version 0.990329
Consensus quality: 73163 bases at least Q40
Consensus quality: 82138 bases at least Q30
Consensus quality: 87186 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 85 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1062: contig of 1062 bp in length
* 1063 1162: gap of unknown length
* 1163 2171: contig of 1009 bp in length
* 2172 2271: gap of unknown length
* 2272 3321: contig of 1050 bp in length
* 3322 3421: gap of unknown length
* 3422 4463: contig of 1042 bp in length
* 4464 4563: gap of unknown length
* 4564 5587: contig of 1024 bp in length
* 5588 5687: gap of unknown length
* 5688 6713: contig of 1025 bp in length
* 6713 6812: gap of unknown length
* 6813 7822: contig of 1010 bp in length
* 7823 7922: gap of unknown length
* 7923 7953: contig of 1831 bp in length
* 7954 9853: gap of unknown length
* 9854 10887: contig of 1034 bp in length

```

```

* 10888 10987: gap of unknown length
* 10988 12401: contig of 1414 bp in length
* 12402 12501: gap of unknown length
* 12502 14037: contig of 1536 bp in length
* 14038 14137: gap of unknown length
* 14138 15197: contig of 1060 bp in length
* 15198 15297: gap of unknown length
* 15298 16644: contig of 1347 bp in length
* 16645 16744: gap of unknown length
* 16745 18177: contig of 1433 bp in length
* 18178 18277: gap of unknown length
* 18278 19658: contig of 1381 bp in length
* 19659 19758: gap of unknown length
* 19759 20806: contig of 1048 bp in length
* 20807 20906: gap of unknown length
* 20907 22312: contig of 1406 bp in length
* 22313 22412: gap of unknown length
* 22413 23922: contig of 1510 bp in length
* 23923 24022: gap of unknown length
* 24023 25433: contig of 1411 bp in length
* 25434 25533: gap of unknown length
* 25534 26809: contig of 1276 bp in length
* 26810 26909: gap of unknown length
* 26910 28404: contig of 1495 bp in length
* 28405 28504: gap of unknown length
* 28505 29670: contig of 1166 bp in length
* 29671 29770: gap of unknown length
* 29771 30994: contig of 1224 bp in length
* 30995 31094: gap of unknown length
* 31095 32675: contig of 1581 bp in length
* 32676 32776: gap of unknown length
* 32776 34130: contig of 1355 bp in length
* 34131 34230: gap of unknown length
* 34231 35489: contig of 1259 bp in length
* 35490 35589: gap of unknown length
* 35590 37026: contig of 1437 bp in length
* 37027 37126: gap of unknown length
* 37127 38796: contig of 1670 bp in length
* 38797 38896: gap of unknown length
* 38897 40001: contig of 1105 bp in length
* 40002 40101: gap of unknown length
* 40102 41115: contig of 1014 bp in length
* 41116 41215: gap of unknown length
* 41216 42569: contig of 1354 bp in length
* 42570 42669: gap of unknown length
* 42670 43830: contig of 1161 bp in length
* 43831 43930: gap of unknown length
* 43931 45281: contig of 1351 bp in length
* 45282 45381: gap of unknown length
* 45382 47441: contig of 2060 bp in length
* 47442 47541: gap of unknown length
* 47542 49026: contig of 1485 bp in length
* 49027 49126: gap of unknown length
* 49127 50263: contig of 1137 bp in length
* 50264 50363: gap of unknown length
* 50364 52441: contig of 2078 bp in length
* 52442 52541: gap of unknown length
* 52542 53981: contig of 1440 bp in length
* 53982 54081: gap of unknown length
* 54082 55502: contig of 1421 bp in length
* 55503 55602: gap of unknown length
* 55603 56893: contig of 1291 bp in length
* 56894 56993: gap of unknown length
* 56994 58953: contig of 1960 bp in length
* 58954 59053: gap of unknown length
* 59054 60351: contig of 1298 bp in length
* 60352 60451: gap of unknown length
* 60452 61848: contig of 1397 bp in length
* 61849 61948: gap of unknown length
* 61949 64132: contig of 2184 bp in length
* 64133 64232: gap of unknown length
* 64233 65624: contig of 1392 bp in length
* 65625 65724: gap of unknown length

Query Match
Best Local Similarity 92.0%; Score 18.4; DB 1; Length 155877;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 TTAGTCTTAGTCTTTAGTT 20
Db 46393 TTAGTCTTAGTCTTAGTT 46412

RESULT 15
AC087327
LOCUS AC087327 160264 bp DNA linear HTG 17-JUL-2001
DEFINITION Trypanosoma brucei chromosome IV clone RPCI93-3112, *** SEQUENCING
IN PROGRESS ***, 1 ordered piece.
ACCESSION AC087327 6 GI:13378168
VERSION AC087327.6
KEYWORDS HTG; HTGS_PHASE2.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei.
REFERENCE 1 (bases 1 to 160264)
AUTHORS El-Sayed,N.M., Ghedin,E., Song,J., Larkin,C., Wanless,D., Jones,K.,
Peterson,D., Hou,L., Zhao,H., Mason,T., Miltzsch,J., Pai,G., Van
Aken,S., Uterback,T., Khalak,H.G., Gerard,C., Leech,V., Ullu,E.,
Melville,S., White,O., Adams,M.D., Donelson,J.E. and Fraser,C.M.
TITLE Trypanosoma brucei CUTa10.1 RPCI93-3112 BAC genomic sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 160264)
AUTHORS El-Sayed,N.M., Khalak,H. and Adams,M.D.
TITLE Direct Submission
JOURNAL Submitted (25-DEC-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
On Mar 19, 2001 this sequence version replaced gi:13277450.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 160264: contig of 160264 bp in length.
FEATURES
location/Qualifiers
1..160264
/organism="Trypanosoma brucei"
/isolate="CUTa10.1"

```

/db\_xref="taxon:5691"  
 /chromosome="IV"  
 /clone="RPC193-3112"  
 BASE COUNT 41037 a 34663 c 39128 g 45436 t  
 ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 160264;  
 Best Local Similarity 95.0%; Pred. No. 39;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Caps 0;  
 OY 1 TTAGTCTTAGTTTAGTT 20  
 ||||| |||||  
 Db 87884 TTAGTTTAGTTTAGTT 87903

Search completed: April 6, 2003, 11:45:30  
 Job time : 428.5 secs



source 1..20  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="Oligonucleotide"

BASE COUNT 3 a 1 c 3 g 13 t

ORIGIN

Query Match 100.0%; Score 20; DB 5; Length 20;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTAGTCTTACTTTTACTT 20  
Db 1 TTAGTCTTACTTTTACTT 20

RESULT 2  
AP004902/c 123078 bp DNA linear PLN 19-JUL-2002  
LOCUS Lotus japonicus genomic DNA, chromosome 2, clone:LTJ04G24, TM0060,  
DEFINITION complete sequence.  
ACCESSION AP004902  
VERSION AP004902.1 GI:21907918  
KEYWORDS HTG.  
SOURCE Lotus japonicus DNA, clone\_lib:LJT library clone:LTJ04G24.  
ORGANISM Lotus japonicus  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteeae;  
Lotus.

REFERENCE 1  
AUTHORS Kaneo,T., Nakamura,Y., Asamizu,E., Kato,T., Sato,S. and Tabata,S.  
TITLE Structural Analysis of a Lotus japonicus Genome. I. Sequence  
Features and Mapping of Sixty-six TAC clones which cover the 6.7 Mb  
Regions of the Genome  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 123078)  
AUTHORS Nakamura,Y.  
TITLE Direct Submission  
JOURNAL Submitted (26-MAR-2002) Yasukazu Nakamura, Kazusa DNA Research  
Institute, Department of Plant Gene Research, 1532-3, Yana,  
Kisarazu, Chiba 292-0812, Japan (E-mail:yn@kazusa.or.jp,  
URL:http://www.kazusa.or.jp, Tel:81-438-52-3935,  
Fax:81-438-52-3934)

FEATURES  
source Location/Qualifiers  
1..123078  
/organism="Lotus japonicus"  
/db\_xref="taxon:34305"  
/chromosome="2"  
/clone="LJT04G24"  
/clone\_lib="LJT library"  
/note="TAC clone:TM0060"

BASE COUNT 40721 a 22328 c 21292 g 38737 t

ORIGIN

Query Match 95.0%; Score 19; DB 7; Length 123078;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TAGTCTTACTTTTACTT 20  
Db 77977 TAGTCTTACTTTTACTT 77959

RESULT 3  
AX057368 20 bp DNA linear PAT 17-JAN-2001  
LOCUS AX057368  
DEFINITION Sequence 2 from Patent WO0075304.  
ACCESSION AX057368  
VERSION AX057368.1 GI:12310109  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct

artificial sequences.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Bachy,M., Sodoyer,R. and Trannoy,E.  
TITLE Immunostimulant oligonucleotide  
JOURNAL Patent: WO 0075304-A 2 14-DEC-2000;  
Aventis Pasteur (FR)

FEATURES  
source Location/Qualifiers  
1..20  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="Oligonucleotide"

BASE COUNT 4 a 0 c 3 g 13 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;  
Best Local Similarity 95.0%; Pred. No. 1.4e+03;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTACTTTTACTT 20  
Db 1 TTAGTATTTACTTTTACTT 20

RESULT 4  
AX057372 20 bp DNA linear PAT 17-JAN-2001  
LOCUS AX057372  
DEFINITION Sequence 6 from Patent WO0075304.  
ACCESSION AX057372  
VERSION AX057372.1 GI:12310113  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct

REFERENCE 1 (bases 1 to 20)  
AUTHORS Bachy,M., Sodoyer,R. and Trannoy,E.  
TITLE Immunostimulant oligonucleotide  
JOURNAL Patent: WO 0075304-A 6 14-DEC-2000;  
Aventis Pasteur (FR)

FEATURES  
source Location/Qualifiers  
1..20  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="Oligonucleotide"

BASE COUNT 3 a 0 c 3 g 14 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;  
Best Local Similarity 95.0%; Pred. No. 1.4e+03;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTACTTTTACTT 20  
Db 1 TTACTTTTACTTTTACTT 20

RESULT 5  
AX057375 20 bp DNA linear PAT 17-JAN-2001  
LOCUS AX057375  
DEFINITION Sequence 9 from Patent WO0075304.  
ACCESSION AX057375  
VERSION AX057375.1 GI:12310116  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct

REFERENCE 1 (bases 1 to 20)  
AUTHORS Bachy,M., Sodoyer,R. and Trannoy,E.  
TITLE Immunostimulant oligonucleotide  
JOURNAL Patent: WO 0075304-A 9 14-DEC-2000;  
Aventis Pasteur (FR)

FEATURES  
source Location/Qualifiers  
1..20  
/organism="synthetic construct"

BASE COUNT 4 a 1 c 3 g 12 t  
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;  
Best Local Similarity 95.0%; Pred. No. 1.4e+03;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTTTACTT 20  
|||||  
Db 1 TTAGTCTTAGTTTACTT 20

RESULT 6  
AX057377 20 bp DNA linear PAT 17-JAN-2001  
LOCUS AX057377  
DEFINITION Sequence 11 from Patent WO075304.  
ACCESSION AX057377  
VERSION AX057377.1 GI:12310118  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Bachy, M., Sodoyer, R. and Tranoy, E.  
TITLE Immunostimulant oligonucleotide  
JOURNAL Patent: WO 0075304-A 11 14-DEC-2000;  
Aventis Pasteur (FR)

FEATURES  
source 1..20  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="Oligonucleotide"

BASE COUNT 3 a 2 c 3 g 12 t  
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;  
Best Local Similarity 95.0%; Pred. No. 1.4e+03;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTTTACTT 20  
|||||  
Db 1 TTAGTCTTAGTTTACTT 20

RESULT 7  
AX057379 20 bp DNA linear PAT 17-JAN-2001  
LOCUS AX057379  
DEFINITION Sequence 13 from Patent WO075304.  
ACCESSION AX057379  
VERSION AX057379.1 GI:12310120  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Bachy, M., Sodoyer, R. and Tranoy, E.  
TITLE Immunostimulant oligonucleotide  
JOURNAL Patent: WO 0075304-A 13 14-DEC-2000;  
Aventis Pasteur (FR)

FEATURES  
source 1..20  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="Oligonucleotide"

BASE COUNT 3 a 1 c 4 g 12 t  
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;  
Best Local Similarity 95.0%; Pred. No. 1.4e+03;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTTTACTT 20  
|||||  
Db 1 TTAGTCTTAGTTTACTT 20

RESULT 8  
AX057380 20 bp DNA linear PAT 17-JAN-2001  
LOCUS AX057380  
DEFINITION Sequence 14 from Patent WO075304.  
ACCESSION AX057380  
VERSION AX057380.1 GI:12310121  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Bachy, M., Sodoyer, R. and Tranoy, E.  
TITLE Immunostimulant oligonucleotide  
JOURNAL Patent: WO 0075304-A 14 14-DEC-2000;  
Aventis Pasteur (FR)

FEATURES  
source 1..20  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="Oligonucleotide"

BASE COUNT 3 a 0 c 4 g 13 t  
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;  
Best Local Similarity 95.0%; Pred. No. 1.4e+03;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTTTACTT 20  
|||||  
Db 1 TTAGTCTTAGTTTACTT 20

RESULT 9  
AX485115/6 65 bp DNA linear PAT 16-AUG-2002  
LOCUS AX485115  
DEFINITION Sequence 2415 from Patent WO02053728.  
ACCESSION AX485115  
VERSION AX485115.1 GI:22319399  
KEYWORDS  
SOURCE Candida albicans.  
ORGANISM Candida albicans.  
REFERENCE 1  
AUTHORS Roemer, T., Jiang, B., Boone, C., Bussey, H. and Ohlsen, K. L.  
TITLE Gene disruption methodologies for drug target discovery  
JOURNAL Patent: WO 02053728-A 2415 11-JUL-2002;  
Eli Lilly Pharmaceuticals, Inc. (US)

FEATURES  
source 1..65  
/organism="Candida albicans"  
/db\_xref="taxon:5476"

BASE COUNT 37 a 6 c 16 g 6 t  
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 65;  
Best Local Similarity 95.0%; Pred. No. 9.5e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTTTACTT 20  
|||||  
Db 46 TTAGTCTTAGTTTACTT 27

RESULT 10  
AF306797/c 2280 bp DNA linear BCT 23-JUL-2001  
LOCUS AF306797  
DEFINITION Vibrio cholerae strain SCE256 toxin co-regulated plus biosynthesis

ACCESSION	protein F (tcpF) and toxin co-regulated pilus virulence regulatory protein (tcpN) genes, complete cds.
VERSION	AF306797.1
KEYWORDS	GI:11139681
SOURCE	Vibrio cholerae.
ORGANISM	Vibrio cholerae
REFERENCE	Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
AUTHORS	1 (bases 1 to 2280)
TITLE	Mukhopadhyay,A.K., Chakraborty,S., Takeda,Y., Nair,G.B. and Berg,D.E.
JOURNAL	Characterization of VPI pathogenicity island and CTXphi prophage in environmental strains of Vibrio cholerae
MEDLINE	J. Bacteriol. 183 (16), 4737-4746 (2001)
PUBMED	21359315
REFERENCE	11466276
AUTHORS	2 (bases 1 to 2280)
TITLE	Mukhopadhyay,A.K., Chakraborty,S., Shimada,T., Takeda,Y., Nair,G.B. and Berg,D.E.
JOURNAL	Direct Submission
FEATURES	Submitted (20-sep-2000) Molecular Microbiology, Washington University Medical School, 660 South Euclid Ave., Saint Louis, MO 63110, USA
SOURCE	Location/Qualifiers
gene	1..2280
CDS	/organism="Vibrio cholerae"
	/strain="SCE256"
	/db_xref="taxon:666"
	189..1199
	/gene="tcpP"
	189..1199
	/gene="tcpF"
	/note="Type III"
	/codon_start=1
	/transl_table=11
	/product="toxin co-regulated pilus biosynthesis protein P"
	/protein_id="AAG31794.1"
	/db_xref="GI:11139682"
	/translation="MKLKLISSTIIILNSALAFNDNSTSYVSTONGDOAGSSGSGTNRYPFACLSGMPSSSLSCYSPPTPSKEGVGPSIEGYTRPEDFPNKRLNSCAVTLESESKYIKPYIGOOTPEMAIPTPOOLTVNLKNGDLXKESKTPYNSGRRLTTDKDGEIVYAKIGNVAEDSPKLPLDIYSKNGLSLDELPSNAVGKLTITKYTHVTVDYSKRDPATFTLLIPDEFERENIDPTKVFEVLGKPSDDGIIRNPKLTTPYYIHDPRPNYSYFIKSDSLHNKGDSDNKYQIKLYTPDSKVSYLKALPIEDKRSQCEGVSD"
gene	1409..2242
CDS	/gene="tcpN"
	1409..2242
	/gene="tcpN"
	/note="Tox"
	/codon_start=1
	/transl_table=11
	/product="toxin co-regulated pilus virulence regulatory protein"
	/protein_id="AAG31795.1"
	/db_xref="GI:11139683"
	/translation="MRENNFSFLNVNYDLKFDTYAEKNVFIDDYKIIFWINKSGSAKLVDKNOISYTIASSSVILLKKNSIOFSLMSLSDSISYCVLTINKFPVNSLRLOODGLIMRINLVNERKKDLLMNCELNDISVIGEIVSYPOLOYSEDPLKIFPSGSPSKYEKKVYNISFITDDLMEKISTCLVKSDITRRNMADICGELRNRIITLKLEBSRCVKRELIILNSIISTISISLMKTGERKIQOIAVOSGFASVSYSTPVKSTMVAPEYLFMGLVAEVE"
BASE COUNT	773 a 334 c 395 g 778 t
ORIGIN	
Query Match	92.0%; Score 18.4; DB 1; Length 2280;
Best Local Similarity	95.0%; Pred. No. 3.3e+02;
Matches 19; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
OY	1 TTAGGTCTTAGTTTTACTT 20                         DB 1308 TTAGTTTTCAGTTTACTT 1289

RESULT 11	AF170083/C	4844 bp	DNA	linear	PLN 20-FEB-2002
LOCUS	AF170083/C				
DEFINITION	Candida albicans putative transcriptional repressor (Ssn6) gene.				
ACCESSION	AF170083				
VERSION	AF170083.2				
KEYWORDS	GI:18767667				
SOURCE	Candida albicans.				
ORGANISM	Candida albicans				
REFERENCE	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; mitosporic Saccharomycetales; Candida.				
AUTHORS	1 (bases 1 to 4844)				
TITLE	Hwang, C.-S. and Kang, S.-O.				
JOURNAL	homologous to Saccharomyces cerevisiae Ssn6				
REFERENCE	2 (bases 1 to 4844)				
AUTHORS	Hwang, C.-S. and Kang, S.-O.				
TITLE	Direct Submission				
JOURNAL	Submitted (19-JUL-1999) Department of Microbiology, College of Natural Science, Seoul National University, Shinlim-Dong, Kwanak-Ku, Seoul 151-742, South Korea				
REFERENCE	3 (bases 1 to 4844)				
AUTHORS	Kang, S.-O.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-FEB-2002) School of Biological Sciences, Institute of Microbiology, Seoul National University, Shinlim-Dong, Kwanak-Ku, Seoul 151-742, South Korea				
REMARK	Sequence update by submitter				
COMMENT	On Feb 20, 2002 this sequence version replaced gi:18000300.				
FEATURES	Location/Qualifiers				
SOURCE	1..4844				
gene	/organism="Candida albicans"				
mrna	/strain="ARCC10231"				
gene	/db.xref="ATCC:10231"				
gene	/db.xref="taxon:5476"				
gene	<1132..>4389				
gene	/gene="Ssn6"				
gene	<1132..>4389				
gene	/gene="Ssn6"				
gene	/product="putative transcriptional repressor"				
gene	1132..4389				
gene	/gene="Ssn6"				
gene	/note="Prf1p, similar to Saccharomyces cerevisiae Ssn6"				
gene	/codon_start=1				
gene	/transl_table=12				
gene	/product="putative transcriptional repressor"				
gene	/protein_id="AAL54912.2"				
gene	/db.xref="GI:18767668"				
gene	/translation="MYTAHTIRKQOQHPPLPLNGCLASGAPRPSHEAAIAQOOQOQOHNGPMTVAALAAASANOQAVARQOQOQOQOQOQOQRLSSALNETTSTWLAIGSLAESLGIERTATASYNLSARHPNNPDLVLAITYRSKDFLAELIYEQALVHNGSEFWGLGHCYLMILDLORAYVAAYQALFYLENPVVKLMHGIGILLIYDRGSLYEAEAFVRLDLIDNPFKANEIYRLGILIKYHOKRLPALECEYILNNPHNPTQEDVMEPFGSVYEOQKDMNGAKADYEKVLIDNPHKAVLIDQLGSLQAASNPPTPANGAAGOHKPRRODITLAKYIQTSLQSLVNDQSNHSMYUGRGVMDRGDTAYAEAFQAVNDARNTFWCSIGVLYTQISOTBDALDATTARLPYISEWVYIDGLTYECNNISDALALADAVARQERLDLPNNPHIKARLEQLTKYQOEBENTPPOPPSSQDPLPQGVNLSLTOOQOQOQOQOQPPPPPOOLQHSQSOQOPPOQPOLOQSLQHSLSLPPQOIQPLQOQAQPLVNOQSPRPHMLNLQPCQOPQOPPLPHLPHTQOQSOIQEKPPTQOOPHNQOPROHOQOQSOQOPRPHQPTQONOSPQLOLPPHSHPRANPHGAQOQPTQGLPDLHNSANITLAPSOVPOQOQOQPHLAPVROBVNHNVPISYAPRPETTLTQILNNPNESTTOVROLKKEKRPPEATVSAAPREAIKYDQDTYIQESAPAAAVSAPASAPQGIKDTIVSTTPATSTTADAVVSVSGEAPVNOVEKAPEDTQIYSQEKVPE				
gene	SQEPVTPAPPAALATAPPEAPDTPDVSVAQVSKATPILQSLVEONTNVSQATKAPESNCKHDLDEKNDSEKILKRPVETTESDVVNOVPYKEKEKEVEVPPRPSSEKREKEVNGSKLKEPLENSKVIDIPOFSNITIAQNEAKSGEETKDTTTPAPQGEVKEVYPPSENYTSKRDVEKDKKDEDEVAADDVQKDDNPPPMKIIDENYIDE"				
BASE COUNT	1722 a 1040 c 736 g 1346 t				
ORIGIN					



```
Query Match          92.0%; Score 18.4; DB 7; Length 4844;
Best Local Similarity 95.0%; Pred. No. 2.6e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTTTAGTT 20
||||| |||||||
Db 392 TTAGTTTTAGTTTAGTT 373

RESULT 12
LOCUS AX252115 5453 bp DNA linear PAT 05-OCT-2001
DEFINITION Sequence 376 from Patent WO0168911.
ACCESSION AX252115
VERSION AX252115.1 GI:15985470
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 5453)
AUTHORS Olek,A., Piepenbrock,C. and Berlin,K.
TITLE Diagnosis of diseases associated with the cell cycle
JOURNAL Patent: WO 0168911-A 376 20-SEP-2001;
Epigenomics AG (DE)
FEATURES
    source Location/Qualifiers
        1..5453
            /organism="synthetic construct"
            /db_xref="taxon:32630"
            /note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 1494 a 101 c 1213 g 2645 t
ORIGIN

Query Match          92.0%; Score 18.4; DB 5; Length 5453;
Best Local Similarity 95.0%; Pred. No. 2.5e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTTTAGTT 20
||||| |||||||
Db 983 TTAGTTTTAGTTTAGTT 1002

RESULT 13
LOCUS AX344477 5453 bp DNA linear PAT 01-FEB-2002
DEFINITION Sequence 324 from Patent WO0200926.
ACCESSION AX344477
VERSION AX344477.1 GI:18492365
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Olek,A., Piepenbrock,C. and Berlin,K.
TITLE Diagnosis of diseases associated with signal transduction
JOURNAL Patent: WO 0200926-A 324 03-JAN-2002;
Epigenomics AG (DE)
FEATURES
    source Location/Qualifiers
        1..5453
            /organism="synthetic construct"
            /db_xref="taxon:32630"
            /note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 1494 a 101 c 1213 g 2645 t
ORIGIN

Query Match          92.0%; Score 18.4; DB 5; Length 5453;
Best Local Similarity 95.0%; Pred. No. 2.5e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTTTAGTT 20
||||| |||||||
Db 983 TTAGTTTTAGTTTAGTT 1002
```

```
RESULT 14
LOCUS AX348882 5453 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 340 from Patent WO0202807.
ACCESSION AX348882
VERSION AX348882.1 GI:18614917
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Olek,A., Piepenbrock,C. and Berlin,K.
TITLE Diagnosis of diseases associated with cell signalling
JOURNAL Patent: WO 0202807-A 340 10-JAN-2002;
Epigenomics AG (DE)
FEATURES
    source Location/Qualifiers
        1..5453
            /organism="synthetic construct"
            /db_xref="taxon:32630"
            /note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 1494 a 101 c 1213 g 2645 t
ORIGIN

Query Match          92.0%; Score 18.4; DB 5; Length 5453;
Best Local Similarity 95.0%; Pred. No. 2.5e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTTTAGTT 20
||||| |||||||
Db 983 TTAGTTTTAGTTTAGTT 1002

RESULT 15
LOCUS AX345350 7238 bp DNA linear PAT 01-FEB-2002
DEFINITION Sequence 421 from Patent WO0200928.
ACCESSION AX345350
VERSION AX345350.1 GI:18493236
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Olek,A., Piepenbrock,C. and Berlin,K.
TITLE Diagnosis of diseases associated with the immune system
JOURNAL Patent: WO 0200928-A 421 03-JAN-2002;
Epigenomics AG (DE)
FEATURES
    source Location/Qualifiers
        1..7238
            /organism="synthetic construct"
            /db_xref="taxon:32630"
            /note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 2414 a 84 c 1401 g 3339 t
ORIGIN

Query Match          92.0%; Score 18.4; DB 5; Length 7238;
Best Local Similarity 95.0%; Pred. No. 2.3e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTTTAGTT 20
||||| |||||||
Db 6279 TTAGTTTTAGTTTAGTT 6298
```

Search completed: April 6, 2003, 11:53:46  
Job time : 268 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 11:21:32 ; Search time 151 Seconds  
(without alignments)  
298.278 Million cell updates/sec

Title: US-09-980-265-10

Perfect score: 20

Sequence: 1 ttactcttagtttagtt 20

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapept 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

N.Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	22	AAF28881 Immunostimulatory
2	19	95.0	20	22	AAF28871 Genetec Immunostim
3	18.4	92.0	20	22	AAF28873 Immunostimulatory
4	18.4	92.0	20	22	AAF28877 Immunostimulatory
5	18.4	92.0	20	22	AAF28880 Immunostimulatory
6	18.4	92.0	20	22	AAF28882 Immunostimulatory
7	18.4	92.0	20	22	AAF28883 Immunostimulatory
8	18.4	92.0	20	22	AAF28885 Immunostimulatory
9	18.4	92.0	5453	24	ABL70450 Chemically treated

10	18.4	92.0	5453	24	AA561413 Human gene regulat
11	18.4	92.0	5453	24	ABK31481 Signal transductio
12	18.4	92.0	7238	24	ABL32448 Human immune syste
13	18.4	92.0	7353	24	ABL32072 Human immune syste
14	18.4	92.0	7353	24	AA28362 Human chemically t
15	18.4	92.0	10682	24	ABN80085 Human chemically m
16	18.4	92.0	11670	24	ABL54325 Chemically treated
17	18.4	92.0	16258	24	ABL70376 Chemically treated
18	18.4	92.0	16258	24	AAK40038 Human chemically p
19	18.4	92.0	18817	24	ABL70162 Chemically treated
20	18.4	92.0	18817	24	ABL34495 Human metastasis a
21	18.4	92.0	21537	24	ABL33998 Human immune syste
22	17.4	87.0	711	24	ABO43350 Oligonucleotide fo
23	17.4	87.0	711	24	ABO43351 Oligonucleotide fo
24	17.4	87.0	885	24	ABO49064 Oligonucleotide fo
25	17.4	87.0	885	24	ABO49065 Oligonucleotide fo
26	17.4	87.0	3745	23	ABL26212 Drosophila melanog
27	17.4	87.0	3775	23	ABL26188 Human immune syste
28	17.4	87.0	5198	24	ABL32279 Human immune syste
29	17.4	87.0	6012	24	ABL70464 Chemically treated
30	17.4	87.0	6012	24	ABK31491 Signal transductio
31	17.4	87.0	7500	24	ABL33115 Human immune syste
32	17.4	87.0	8467	24	ABL32108 Human immune syste
33	17.4	87.0	15698	24	ABL34141 Human immune syste
34	17.4	87.0	21667	23	ABL03568 Drosophila melanog
35	16.8	84.0	20	22	AAF28872 Immunostimulatory
36	16.8	84.0	20	22	AAF28874 Immunostimulatory
37	16.8	84.0	20	22	AAF28875 Immunostimulatory
38	16.8	84.0	20	22	AAF28876 Immunostimulatory
39	16.8	84.0	20	22	AAF28878 Immunostimulatory
40	16.8	84.0	20	22	AAF28879 Immunostimulatory
41	16.8	84.0	20	22	AAF28884 Immunostimulatory
42	16.8	84.0	20	22	AAF28886 Immunostimulatory
43	16.8	84.0	20	22	AAF28887 Immunostimulatory
44	16.8	84.0	488	24	ABL93059 Rat metastatic tum
45	16.8	84.0	569	24	ABO48528 Oligonucleotide fo

#### ALIGNMENTS

RESULT 1	
AAF28881	AAF28881 standard; DNA; 20 BP.
ID	
AC	AAF28881;
XX	
DT	09-MAY-2001 (first entry)
XX	
DE	Immunostimulatory oligonucleotide #10 as vaccine adjuvant.
XX	
KW	Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KW	prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW	cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KW	phosphorothioate; ss.
XX	
OS	Synthetic.
XX	
FH	Key
FT	modified_base 1..20
FT	Location/Qualifiers
FT	/*tag= a
FT	/note= "contain phosphorothioate internucleotide linkages"
XX	
PN	WO200075304-A1.
XX	
PD	14-DEC-2000.
XX	
PF	08-JUN-2000; 2000WO-FR01566.
XX	
PR	08-JUN-1999; 99FR-0007457.
PR	06-AUG-1999; 99FR-0010378.
XX	

PA (AVET ) AVENTIS PASTEUR.  
 XX  
 PI Bachy M, Sodoyer R, Tranoy E;  
 XX WPI; 2001-041317/05.  
 DR  
 XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in  
 PT vaccines for human use, induce lymphocyte proliferation and cytokine  
 PT secretion -  
 XX  
 PS Example 4; Page 13; 30pp; French.  
 CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an  
 CC immunostimulatory oligonucleotide of the invention which contains at  
 CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The  
 CC oligonucleotides do not contain any CG dinucleotides in which C are  
 CC unethyalted. The oligonucleotides are used as human immunostimulants  
 CC and as adjuvants in therapeutic and prophylactic vaccines for human use.  
 CC They induce proliferation of human lymphocytes, induce secretion of  
 CC cytokines, especially interleukin-10 or interferon-gamma and increase  
 CC expression of the CD86 activation marker or the CD25 cytokine receptor  
 CC on human B lymphocytes.  
 XX  
 SQ Sequence 20 BP; 3 A; 1 C; 3 G; 13 T; 0 other;  
 Query Match 100.0%; Score 20; DB 22; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 33;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 TTACTCTTAGTTTGTAGTT 20  
 Db 1 TTACTCTTAGTTTGTAGTT 20  
 RESULT 2  
 AAF28871  
 ID AAF28871 standard; DNA: 20 BP.  
 AC AAF28871;  
 XX  
 DT 09-MAY-2001 (first entry)  
 XX  
 DE Generic immunostimulatory oligonucleotide as vaccine adjuvant.  
 XX  
 KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;  
 KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;  
 KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor; ss.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200075304-A1.  
 XX  
 PD 14-DEC-2000.  
 XX  
 PF 08-JUN-2000; 2000WO-FR01566.  
 XX  
 PR 08-JUN-1999; 99FR-0007457.  
 PR 06-AUG-1999; 99FR-0010378.  
 XX  
 PA (AVET ) AVENTIS PASTEUR.  
 XX  
 PI Bachy M, Sodoyer R, Tranoy E;  
 XX  
 DR WPI; 2001-041317/05.  
 XX  
 PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in  
 PT vaccines for human use, induce lymphocyte proliferation and cytokine  
 PT secretion -  
 XX  
 PS Claim 8; Page 17; 30pp; French.  
 CC This sequence represents a generic example of an immunostimulatory  
 CC oligonucleotide of the invention which contains at least one sequence

CC 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The oligonucleotides  
 CC do not contain any CG dinucleotides in which C are unethyalted.  
 CC The oligonucleotides are used as human immunostimulants and as adjuvants  
 CC in therapeutic and prophylactic vaccines for human use. They induce  
 CC proliferation of human lymphocytes, induce secretion of cytokines,  
 CC especially interleukin-10 or interferon-gamma and increase expression  
 CC of the CD86 activation marker or the CD25 cytokine receptor on human B  
 CC lymphocytes.  
 XX  
 SQ Sequence 20 BP; 3 A; 1 C; 3 G; 12 T; 1 other;  
 Query Match 95.0%; Score 19; DB 22; Length 20;  
 Best Local Similarity 95.0%; Pred. No. 82;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 TTACTCTTAGTTTGTAGTT 20  
 Db 1 TTACTCTTAGTTTGTAGTT 20  
 RESULT 3  
 AAF28873  
 ID AAF28873 standard; DNA: 20 BP.  
 AC AAF28873;  
 XX  
 DT 09-MAY-2001 (first entry)  
 XX  
 DE Immunostimulatory oligonucleotide #2 as vaccine adjuvant.  
 XX  
 KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;  
 KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;  
 KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;  
 KW phosphorothioate; ss.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200075304-A1.  
 XX  
 PD 14-DEC-2000.  
 XX  
 PF 08-JUN-2000; 2000WO-FR01566.  
 XX  
 PR 08-JUN-1999; 99FR-0007457.  
 PR 06-AUG-1999; 99FR-0010378.  
 XX  
 PA (AVET ) AVENTIS PASTEUR.  
 XX  
 PI Bachy M, Sodoyer R, Tranoy E;  
 XX  
 DR WPI; 2001-041317/05.  
 XX  
 PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in  
 PT vaccines for human use, induce lymphocyte proliferation and cytokine  
 PT secretion -  
 XX  
 PS Example 4; Page 13; 30pp; French.  
 CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an  
 CC immunostimulatory oligonucleotide of the invention which contains at  
 CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The  
 CC oligonucleotides do not contain any CG dinucleotides in which C are  
 CC unethyalted. The oligonucleotides are used as human immunostimulants  
 CC and as adjuvants in therapeutic and prophylactic vaccines for human use.  
 CC They induce proliferation of human lymphocytes, induce secretion of  
 CC cytokines, especially interleukin-10 or interferon-gamma and increase  
 CC expression of the CD86 activation marker or the CD25 cytokine receptor

CC on human B lymphocytes.

XX Sequence 20 BP; 4 A; 0 C; 3 G; 13 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 20;

Best Local Similarity 95.0%; Pred. No. 1.4e+02;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTTTACTT 20

DB 1 TTAGTCTTAGTTTACTT 20

RESULT 4

AAF28877

ID AAF28877 standard; DNA; 20 BP.

XX AAF28877;

DT 09-MAY-2001 (first entry)

DE Immunostimulatory oligonucleotide #6 as vaccine adjuvant.

XX Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;

KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;

KM cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;

XX phosphorothioate; ss.

OS Synthetic.

XX Key Location/Qualifiers

FT modified\_base 1..20

FT /\*tag= a

FT /note= "contain phosphorothioate internucleotide linkages"

XX WO200075304-A1.

XX 14-DEC-2000.

XX 08-JUN-2000; 2000WO-FR01566.

XX 08-JUN-1999; 99FR-0007457.

XX 06-AUG-1999; 99FR-0010378.

XX (AVET ) AVENTIS PASTEUR.

XX Bachy M, Sodoyer R, Tranoy E;

XX WPI; 2001-041317/05.

XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in

PT vaccines for human use, induce lymphocyte proliferation and cytokine

PS secretion -

XX Example 4; Page 13; 30pp; French.

XX Oligonucleotides AAF28872-AAF28887 represent specific examples of an

CC immunostimulatory oligonucleotide of the invention which contains at

CC least one sequence 5'-TTNNN2TT-3' where N1 and N2 are A, T, C or G. The

CC oligonucleotides do not contain any CG dinucleotides in which C are

CC unethyalted. The oligonucleotides are used as human immunostimulants

CC and as adjuvants in therapeutic and prophylactic vaccines for human use.

CC They induce proliferation of human lymphocytes, induce secretion of

CC cytokines, especially interleukin-10 or interferon-gamma and increase

CC expression of the CD86 activation marker or the CD25 cytokine receptor

CC on human B lymphocytes.

XX Sequence 20 BP; 3 A; 0 C; 3 G; 14 T; 0 other;

XX Query Match 92.0%; Score 18.4; DB 22; Length 20;

XX Best Local Similarity 95.0%; Pred. No. 1.4e+02;

XX Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTTTACTT 20

DB 1 TTAGTCTTAGTTTACTT 20

RESULT 5

AAF28880

ID AAF28880 standard; DNA; 20 BP.

XX AAF28880;

DT 09-MAY-2001 (first entry)

DE Immunostimulatory oligonucleotide #9 as vaccine adjuvant.

XX Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;

KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;

KM cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;

XX phosphorothioate; ss.

OS Synthetic.

XX Key Location/Qualifiers

FT modified\_base 1..20

FT /\*tag= a

FT /note= "contain phosphorothioate internucleotide linkages"

XX WO200075304-A1.

XX 14-DEC-2000.

XX 08-JUN-2000; 2000WO-FR01566.

XX 08-JUN-1999; 99FR-0007457.

XX 06-AUG-1999; 99FR-0010378.

XX (AVET ) AVENTIS PASTEUR.

XX Bachy M, Sodoyer R, Tranoy E;

XX WPI; 2001-041317/05.

XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in

PT vaccines for human use, induce lymphocyte proliferation and cytokine

PS secretion -

XX Example 4; Page 13; 30pp; French.

XX Oligonucleotides AAF28872-AAF28887 represent specific examples of an

CC immunostimulatory oligonucleotide of the invention which contains at

CC least one sequence 5'-TTNNN2TT-3' where N1 and N2 are A, T, C or G. The

CC oligonucleotides do not contain any CG dinucleotides in which C are

CC unethyalted. The oligonucleotides are used as human immunostimulants

CC and as adjuvants in therapeutic and prophylactic vaccines for human use.

CC They induce proliferation of human lymphocytes, induce secretion of

CC cytokines, especially interleukin-10 or interferon-gamma and increase

CC expression of the CD86 activation marker or the CD25 cytokine receptor

CC on human B lymphocytes.

XX Sequence 20 BP; 4 A; 1 C; 3 G; 12 T; 0 other;

XX Query Match 92.0%; Score 18.4; DB 22; Length 20;

XX Best Local Similarity 95.0%; Pred. No. 1.4e+02;

XX Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX OY 1 TTAGTCTTAGTTTACTT 20

XX DB 1 TTAGTCTTAGTTTACTT 20

XX RESULT 6

```
AAF28882
ID AAF28882 standard; DNA; 20 BP.
XX
AC AAF28882;
XX
DT 09-MAY-2001 (first entry)
XX
DE Immunostimulatory oligonucleotide #11 as vaccine adjuvant.
XX
KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KW phosphorothioate; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1..20
FT /tag= a
FT /note= "contain phosphorothioate internucleotide
FT linkages"
XX
PN WO200075304-A1.
XX
PD 14-DEC-2000.
XX
PF 08-JUN-2000; 2000WO-FR01566.
XX
PR 08-JUN-1999; 99FR-0007457.
PR 06-AUG-1999; 99FR-0010378.
XX
PA (AVET ) AVENTIS PASTEUR.
XX
PI Bachy M, Sodoyer R, Tranney E;
XX
DR WPI; 2001-041317/05.
XX
PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
PT vaccines for human use, induce lymphocyte proliferation and cytokine
PT secretion -
XX
PS Example 4; Page 13; 30pp; French.
XX
CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an
CC immunostimulatory oligonucleotide of the invention which contains at
CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The
CC oligonucleotides do not contain any CG dinucleotides in which C are
CC unmethylated. The oligonucleotides are used as human immunostimulants
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
CC They induce proliferation of human lymphocytes, induce secretion of
CC cytokines, especially interleukin-10 or interferon-gamma and increase
CC expression of the CD86 activation marker or the CD25 cytokine receptor
CC on human B lymphocytes.
XX
SQ Sequence 20 BP; 3 A; 2 C; 3 G; 12 T; 0 other;
XX
Query Match 92.0%; Score 18.4; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTAGTCTTAGTTTTCAGTT 20
DB 1 TTAGTCTTAGTTTTCAGTT 20
XX
RESULT 7
AAF28883
ID AAF28883 standard; DNA; 20 BP.
XX
AC AAF28883;
XX
DT 09-MAY-2001 (first entry)
XX
```

```
DE Immunostimulatory oligonucleotide #12 as vaccine adjuvant.
XX
KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KW phosphorothioate; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1..20
FT /tag= a
FT /note= "contain phosphorothioate internucleotide
FT linkages"
XX
PN WO200075304-A1.
XX
PD 14-DEC-2000.
XX
PF 08-JUN-2000; 2000WO-FR01566.
XX
PR 08-JUN-1999; 99FR-0007457.
PR 06-AUG-1999; 99FR-0010378.
XX
PA (AVET ) AVENTIS PASTEUR.
XX
PI Bachy M, Sodoyer R, Tranney E;
XX
DR WPI; 2001-041317/05.
XX
PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
PT vaccines for human use, induce lymphocyte proliferation and cytokine
PT secretion -
XX
PS Example 4; Page 13; 30pp; French.
XX
CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an
CC immunostimulatory oligonucleotide of the invention which contains at
CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The
CC oligonucleotides do not contain any CG dinucleotides in which C are
CC unmethylated. The oligonucleotides are used as human immunostimulants
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
CC They induce proliferation of human lymphocytes, induce secretion of
CC cytokines, especially interleukin-10 or interferon-gamma and increase
CC expression of the CD86 activation marker or the CD25 cytokine receptor
CC on human B lymphocytes.
XX
SQ Sequence 20 BP; 3 A; 1 C; 4 G; 12 T; 0 other;
XX
Query Match 92.0%; Score 18.4; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTAGTCTTAGTTTTCAGTT 20
DB 1 TTAGTCTTAGTTTTCAGTT 20
XX
RESULT 8
AAF28885
ID AAF28885 standard; DNA; 20 BP.
XX
AC AAF28885;
XX
DT 09-MAY-2001 (first entry)
XX
DE Immunostimulatory oligonucleotide #14 as vaccine adjuvant.
XX
KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KW phosphorothioate; ss.
XX
```

OS	Synthetic.	Location/Qualifiers
XX	Key	1..20
EH	modified_base	/*tag="a
FT		/note="contain phosphorothioate internucleotide
FT		linkages"
XX		
XX	WO200075304-A1.	
XX		
XX	14-DEC-2000.	
PD		
XX		
PF	08-JUN-2000; 2000WO-FR01566.	
XX		
XX	08-JUN-1999; 99FR-0007457.	
PR	06-AUG-1999; 99FR-0010378.	
XX		
PA	(AVET ) AVENTIS PASTEUR.	
XX		
PI	Bachy M, Sodoayer R, Trannoy E;	
XX		
DR	WPI; 2001-041317/05.	
XX		
PT	New immunostimulatory oligonucleotides, useful e.g. as adjuvants in	
PT	vaccines for human use, induce lymphocyte proliferation and cytokine	
PT	secretion -	
XX		
PS	Example 4; Page 13; 30pp; French.	
XX		
CC	Oligonucleotides AAF28872-AAF28887 represent specific examples of an	
CC	immunostimulatory oligonucleotide of the invention which contains at	
CC	least one sequence 5'-TTNIN2T-3' where N1 and N2 are A, T, C or G. The	
CC	oligonucleotides do not contain any CG dinucleotides in which C are	
CC	unmethylated. The oligonucleotides are used as human immunostimulants	
CC	and as adjuvants in therapeutic and prophylactic vaccines for human use.	
CC	They induce proliferation of human lymphocytes, induce secretion of	
CC	cytokines, especially interleukin-10 or interferon-gamma and increase	
CC	expression of the CD86 activation marker or the CD25 cytokine receptor	
CC	on human B lymphocytes.	
XX		
XX		
SQ	Sequence 20 BP; 3 A; 0 C; 4 G; 13 T; 0 other;	
	Query Match	92.0%; Score 18.4; DB 22; Length 20;
	Best Local Similarity	95.0%; Pred. No. 1.4e+02;
	Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
OY	1 TTAGTCTTACTTTTACTT 20	
DB	1 TTAGTTGTAGTTTTTACTT 20	
RESULT 9		
ABL70450	ABL70450 standard; DNA; 5453 BP.	
XX		
AC	ABL70450;	
XX		
DE	01-JUL-2002 (first entry)	
XX		
XX		
DE	Chemically treated cell signalling DNA sequence complementary to#170.	
XX		
KW	Cell signalling; cytosine methylation; cell signalling disease;	
KM	cancer; tumour; cytostatic; ds.	
XX		
OS	Unidentified.	
XX		
XX	WO200202807-A2.	
PN		
PD	10-JAN-2002.	
XX		
XX	29-JUN-2001; 2001WO-EP07471.	
PF		
XX		
XX	30-JUN-2000; 2000DE-1032529.	
PR		

PR	01-SEP-2000:	2000DE-1043826.
XX		
PA	(EPig-) EPiGENOMICS AG.	
XX		
PI	Olek A, Piepenbrock C, Berlin K:	
XX		
DR	WPI; 2002-154758/20.	
XX		
PT	Nucleic acid, useful for diagnosis and therapy of diseases associated	
XX	with cell signalling e.g. cancer, comprises chemically modified genomic	
PT	sequences of genes associated with cell signalling	
XX		
PS	Claim 1; SEQ ID NO 340; 24pp+sequence listing; English.	
XX		
CC	The invention relates to a nucleic acid comprising a sequence of at least	
CC	18 bases of a segment of chemically pretreated DNA of genes associated	
CC	with cell signalling. The activity of the modified sequences of the	
CC	invention may be described as cytostatic. The object of the invention is	
CC	to provide the chemically modified DNA of genes associated with cell	
CC	signalling, as well as oligonucleotides and/or PNA-oligomers for	
CC	detecting cytosine methylations, as well as a method which is	
CC	particularly suitable for the diagnosis and/or therapy of genetic and	
CC	epigenetic parameters of genes associated with cell signalling. The	
CC	chemically modified DNA provided by the invention is useful for diagnosis	
CC	and therapy of diseases such as solid tumours and cancer. The sequences	
CC	given in records AB1/0111-AB1/0626 represent chemically pre-treated	
CC	genomic DNA's of genes associated with cell signalling.	
CC	Note: The sequence data for this patent is not represented in the printed	
CC	specification, but is based on sequence information supplied by the	
CC	European Patent Office.	
XX		
SQ	Sequence 5453 BP; 1494 A; 101 C; 1213 G; 2645 T; 0 other;	
XX		
Query Match	92.0%: Score 18.4; DB 24; Length 5453;	
Best Local Similarity	95.0%: Pred No. 1.2e+02;	
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0		
QY	1 TTAGTCTCTAGTTTATGTT 20	
Db	983 TTAGCTTTTGTAGTTTATGTT 1002	
XX		
RESULT 10		
ID	AAS61413 standard; DNA; 5453 BP.	
XX		
AC	AAS61413;	
XX		
DT	29-JAN-2002 (first entry)	
XX		
DE	Human gene regulation-associated gene oligonucleotide #368.	
XX		
KW	Human: Gene regulation-associated gene; severe combined immunodeficiency;	
KW	cardiac damage; inflammatory response; Haemophilia; Werner syndrome;	
KW	asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;	
KW	renal disease; Preeclampsia; cardiac allograft vascular disease;	
KW	colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;	
KW	immunostimulant; cardiac; antiinflammatory; coagulant; antiasthmatic;	
KW	nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.	
XX		
OS	Homo sapiens.	
XX		
PN	WO2001/7375-A2.	
XX		
PD	18-OCT-2001.	
XX		
PF	06-APR-2001; 2001WO-EP033968.	
XX		
PR	06-APR-2000; 2000DE-1019058.	
PR	07-APR-2000; 2000DE-1019173.	
PR	30-JUN-2000; 2000DE-1032529.	
XX		
PR	01-SEP-2000; 2000DE-1043826.	
XX		

PA (EPIC-) EPIGENOMICS AG.  
XX  
PI Olek A, Piepenbrock C, Berlin K;  
XX  
DR WPI; 2002-017470/02.  
XX  
PT New nucleic acid sequences from chemically modified genes associated  
PT with gene regulation, useful for analysing cytosine methylations for  
PT diagnosis and therapy of diseases e.g. severe combined immunodeficiency  
XX  
XX  
PS Disclosure; SEQ ID No 376; 26pp; English.  
XX  
CC The invention relates to 224 nucleic acid sequences comprising at least  
CC 18 bases of a chemically pretreated gene associated with gene regulation  
CC selected from 43 known genes (or complementary sequences). The  
CC chemical pretreatment converts cytosine bases unmethylated at the  
CC 5-position to uracil or another base with hybridisation behaviour  
CC dissimilar to cytosine, to enable analysis of cytosine methylations.  
CC The DNA sequences, oligomers (or sets/arrays) and method are  
CC useful in the diagnosis of diseases (or predisposition to diseases)  
CC associated with gene regulation and in therapy of such diseases, by  
CC enabling analysis of the cytosine methylation patterns of such genes,  
CC kits are provided. They are especially useful in diagnosis  
CC and therapy of e.g. severe combined immunodeficiency disease, cardiac  
CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,  
CC asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,  
CC pre-eclampsia, graft versus-host disease. The present sequence is a  
CC sequence included in the sequence data for this specification and is  
CC associated with the human gene regulation-associated genes.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 5453 BP; 1494 A; 101 C; 1213 G; 2645 T; 0 other;  
XX  
Query Match 92.0%; Score 18.4; DB 24; Length 5453;  
Best Local Similarity 95.0%; Pred. NO. 1.2e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Oy 1 TTAGTCTTAGTTTGGT 20  
Db 983 TTAGTTTGTAGTTTGGT 1002  
XX  
RESULT 11  
ABK31481  
ID ABK31481 standard; DNA: 5453 BP.  
XX  
AC ABK31481;  
XX  
DT 23-APR-2002 (first entry)  
XX  
DE Signal transduction associated gene modified complementary DNA #162.  
XX  
KW Human; signal transduction associated gene; cytosine methylation state;  
KW CpG island; signal transduction associated disease; solid tumour; cancer;  
KW antitumour; cytostatic; mutant; ds.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN WO200200926-A2.  
XX  
PD 03-JAN-2002.  
XX  
PF 29-JUN-2001; 2001WO-EP07472.  
XX  
PR 30-JUN-2000; 2000DE-1032529.  
PR 01-SEP-2000; 2000DE-1043826.  
XX  
PA (EPIC-) EPIGENOMICS AG.

XX  
XX Olek A, Piepenbrock C, Berlin K;  
XX  
DR WPI; 2002-147896/19.  
XX  
PT Oligonucleotide for diagnosis and therapy of diseases associated with  
PT signal transduction e.g. cancer, comprises chemically modified genomic  
PT sequences of genes associated with signal transduction  
XX  
XX  
PS Claim 1; SEQ ID No 324; 24pp; English.  
XX  
CC The present invention relates to chemically modified DNA sequences of  
CC signal transduction associated genes. The DNA sequences are chemically  
CC modified using a solution of bisulphite, hydrogen sulphite or  
CC disulphite. Also disclosed are oligonucleotides and/or PNA oligomers  
CC for detecting the cytosine methylation state (CpG islands) of these  
CC genes, and a method for the diagnosis and/or therapy of genetic and  
CC epigenetic parameters of genes associated with signal transduction.  
CC The genomic DNA can be obtained from cells or cellular components which  
CC contain DNA, e.g. cell lines, biopsies, blood, sputum, stool, urine,  
CC cerebral-spinal fluid, tissue embedded in paraffin such as tissue from  
CC eyes, intestine, kidney, brain, heart, prostate, lung, breast or liver,  
CC histologic object slides, and all their possible combinations. The  
CC sequences of the invention are useful for the diagnosis and therapy of  
CC diseases associated with signal transduction e.g. solid tumours and  
CC cancer. ABK31158-ABK31545 represent chemically pretreated genomic DNA  
CC sequences of different genes associated with signal transduction, or  
CC their complementary sequences.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from the  
CC European Patent Office.  
XX  
SQ Sequence 5453 BP; 1494 A; 101 C; 1213 G; 2645 T; 0 other;  
XX  
Query Match 92.0%; Score 18.4; DB 24; Length 5453;  
Best Local Similarity 95.0%; Pred. NO. 1.2e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Oy 1 TTAGTCTTAGTTTGGT 20  
Db 983 TTAGTTTGTAGTTTGGT 1002  
XX  
RESULT 12  
ABL32448  
ID ABL32448 standard; DNA: 7238 BP.  
XX  
AC ABL32448;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Human immune system associated gene SEQ ID NO: 421.  
XX  
KW Human immune system disease; cytosine methylation; antiasthmatic;  
KW antiarteriosclerotic; antianaemic; cycostatic; nootropic;  
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
KW antilemmatic; antiarthritic; antidiabetic; antipsoriatic;  
KW antilemmatoly; cancer; eye disease; arteriosclerosis; anaemia;  
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
KW gene; ds.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN WO200200928-A2.  
XX  
PD 03-JAN-2002.  
XX  
PF 02-JUL-2001; 2001WO-EP07537.  
XX  
PR 30-JUN-2000; 2000DE-1032529.  
PR 01-SEP-2000; 2000DE-1043826.  
XX



PA (EPiG-) EPIGENOMICS AG.  
XX  
XX Olek A. Piepenbrock C, Berlin K;  
XX  
XX WPI: 2002-130909/17.  
XX  
XX Nucleic acid comprising fragment of chemically modified gene, useful  
PT for diagnosis and treatment of diseases associated with abnormal  
PT cytosine methylation  
XX  
PS Claim 1; SEQ ID NO 421; 32pp + Sequence Listing; German.  
XX  
XX The present invention provides a number of human immune system associated  
CC genes which are modified by the methylation of cytosines. The sequences  
CC can be used in the diagnosis and treatment of immune system disorders,  
CC including eye diseases such as retinopathy, neovascular glaucoma and  
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
CC diseases. The present sequence is a gene of the invention.  
XX  
SQ Sequence 7238 BP; 2414 A; 84 C; 1401 G; 3339 T; 0 other;  
XX  
Query Match 92.0%; Score 18.4; DB 24; Length 7238;  
Best Local Similarity 95.0%; Pred. No. 1.2e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 TTAGTCTTAGTTTATTAGTT 20  
|||||  
Db 6279 TTAGTTTATTAGTTTATTAGTT 6298  
RESULT 13  
ABL32072  
ID ABL32072 standard; DNA; 7353 BP.  
XX  
AC ABL32072;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Human immune system associated gene SEQ ID NO: 45.  
XX  
XX Human; immune system disease; cytosine methylation; antiasthmatic;  
KW antiarteriosclerotic; antihaemic; cytosatic; nootropic;  
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
KW antirheumatic; antiarthritic; antidiabetic; antiparasitic;  
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
KW gene; ds.  
XX  
XX Homo sapiens.  
OS  
XX WO200200928-A2.  
XX  
XX 03-JAN-2002.  
XX  
XX 02-JUL-2001; 2001WO-EP07537.  
XX  
XX 30-JUN-2000; 2000DE-1032529.  
PR 01-SEP-2000; 2000DE-1043826.  
XX  
XX (EPiG-) EPIGENOMICS AG.  
XX  
XX Olek A. Piepenbrock C, Berlin K;  
XX  
XX WPI: 2002-130909/17.  
XX  
XX Nucleic acid comprising fragment of chemically modified gene, useful  
PT for diagnosis and treatment of diseases associated with abnormal  
PT cytosine methylation  
XX  
XX Claim 1; SEQ ID NO 45; 32pp + Sequence Listing; German.

XX  
XX The present invention provides a number of human immune system associated  
CC genes which are modified by the methylation of cytosines. The sequences  
CC can be used in the diagnosis and treatment of immune system disorders,  
CC including eye diseases such as retinopathy, neovascular glaucoma and  
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
CC diseases. The present sequence is a gene of the invention.  
XX  
SQ Sequence 7353 BP; 1544 A; 384 C; 1956 G; 3469 T; 0 other;  
XX  
Query Match 92.0%; Score 18.4; DB 24; Length 7353;  
Best Local Similarity 95.0%; Pred. No. 1.2e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 TTAGTCTTAGTTTATTAGTT 20  
|||||  
Db 2404 TTAGTTTATTAGTTTATTAGTT 2423  
RESULT 14  
AAD28362  
ID AAD28362 standard; DNA; 7353 BP.  
XX  
XX AAD28362;  
XX  
XX 22-APR-2002 (first entry)  
XX  
DE Human chemically treated genomic DNA #3.  
XX  
XX Human; cytostatic; antidepressant; neuroleptic; nootropic; antiaddictive;  
KW adrenergic alpha-1C-receptor; cytosine methylation; therapy; alcoholism;  
KW behavioural disorder; neurological; psychiatric; cancer; schizophrenia;  
KW Tourette's syndrome; smoking; human immunodeficiency virus dementia;  
KW drug abuse; migraine; ds.  
XX  
XX Homo sapiens.  
OS  
XX WO200202809-A2.  
XX  
XX 10-JAN-2002.  
XX  
XX 02-JUL-2001; 2001WO-EP07540.  
XX  
XX 30-JUN-2000; 2000DE-1032529.  
PR 01-SEP-2000; 2000DE-1043826.  
XX  
XX (EPiG-) EPIGENOMICS AG.  
XX  
XX Olek A. Piepenbrock C, Berlin K;  
XX  
XX WPI: 2002-154759/20.  
XX  
XX Novel nucleic acid useful for diagnosis and therapy of behavioral  
PT disorder, neurological disorder and cancer, comprises a sequence of a  
PT segment of chemically pretreated DNA of adrenergic alpha-1C-receptor  
PT gene  
XX  
XX Claim 1; Page 32-36; 190pp; English.  
XX  
XX The invention relates to nucleic acids comprising a segment of chemically  
CC pretreated DNA of adrenergic alpha-1C-receptor gene. The invention also  
CC relates to oligonucleotides or peptide nucleic acid (PNA) oligomers  
CC useful for detecting cytosine methylations. The pretreated DNA is useful  
CC for the diagnosis or therapy of behavioural disorders, neurological  
CC disorders and cancer, in particular major depressive disorder, Tourette's  
CC syndrome, schizophrenia, psychiatric and neurological disorders, smoking,  
CC drug abuse, alcoholism, personality traits, compulsive gambling, human  
CC immunodeficiency virus dementia, migraine, behaviours in schizophrenic  
CC and schizoaffective patients, and suicidal behaviour in patients with  
CC schizophrenia. The nucleic acid is useful for detecting the methylation  
CC state of all Cpg dinucleotides and/or single nucleotide polymorphisms

```
CC (SNPs). The present sequence is human chemically treated genomic DNA.
XX
SQ Sequence 7353 BP; 1544 A; 384 C; 1956 G; 3469 T; 0 other;
Query Match 92.0%; Score 18.4; DB 24; Length 7353;
Best Local Similarity 95.0%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TTAGTCTTAGTCTTTAGTT 20
   ||||| ||||| |||||
Db 2404 TTACTTATAGTTTGTAGTT 2423

RESULT 15
ABN80085
ID ABN80085 standard; DNA; 10682 BP.
XX
AC ABN80085;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human chemically modified disease associated gene SEQ ID NO 102.
XX
KW Human; development; homeobox gene; HOX; diabetes; cancer; apoptosis;
KW heart disease; epilepsy; histone deacetylation; muscular dystrophy;
KW dwarfism; single nucleotide polymorphism; SNP; cytosine methylation;
KW antidiabetic; cytostatic; anticonvulsant; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200200927-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP07536.
XX
PR 30-JUN-2000; 2000DE-1032529.
XX
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPiG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130908/17.
XX
PT Novel nucleic acid useful for diagnosis and therapy of diseases
PT associated with development genes such as diabetes, comprises a
PT sequence of a segment of chemically pretreated DNA of genes associated
PT with development.
XX
PS Claim 1; SEQ ID NO 102; 27pp; English.
XX
CC The invention relates to a nucleic acid (I) comprising a sequence at
CC least 18 bases in length of a segment of chemically pretreated DNA (II)
CC of genes associated with development selected from 87 genes listed in
CC the specification such as ACCPN, ADFN, or AFDI and comprising one of 350
CC sequences (ABN79984-ABN80333) or their complements. The invention is
CC useful for the diagnosis or therapy of diseases associated with
CC development genes, in particular disease related to homeobox containing
CC genes (HOX), like diabetes, cancer, apoptosis related diseases, syndromes
CC associated with congenital heart disease, epilepsy, diseases related to
CC histone deacetylation, Curranio syndrome, diseases related with the
CC development of the brain and limb girdle muscular dystrophy and dwarfism.
CC Oligomers specific to each of the genes are useful for detecting the
CC methylation state of all CpG dinucleotides within the 350 sequences or
CC (II) and their complementary sequences, as primer oligonucleotides for
CC the amplification of the 350 sequences, (II) and/or their complements and
CC as oligomer probes for detecting the cytosine methylation state and/or
CC single nucleotide polymorphisms (SNPs).
CC Note: The sequence data for this patent did not form part of the printed
CC specification but is based on sequence information supplied to Derwent by
CC the European Patent Office.
```

```
XX
SQ Sequence 10682 BP; 3158 A; 263 C; 2318 G; 4943 T; 0 other;
Query Match 92.0%; Score 18.4; DB 24; Length 10682;
Best Local Similarity 95.0%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TTAGTCTTAGTCTTTAGTT 20
   ||||| ||||| |||||
Db 5298 TTACTTTTATAGTTTGTAGTT 5317

Search completed: April 6, 2003, 12:05:26
Job time : 155 secs
```

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 11:12:37 ; Search time 45.25 Seconds  
(without alignments)  
135.548 Million cell updates/sec

Title: US-09-980-265-10

Perfect score: 20

Sequence: 1 ttagtcttagtttagtt 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued\_Patents\_NA.\*  
2: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*  
7: /cgn2\_6/ptodata/2/ina/Backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description
	1	16.8	84.0	1602	2		US-08-770-544-3	Sequence 3, Appl
C	2	15.8	79.0	362	4		US-09-605-785-442	Sequence 442, App
C	3	15.8	79.0	362	4		US-09-439-313-442	Sequence 442, App
C	4	15.8	79.0	362	4		US-09-352-616A-442	Sequence 442, App
5	15.8	79.0	427	4			US-09-556-877-46	Sequence 46, App
6	15.8	79.0	427	4			US-09-620-412C-46	Sequence 46, App
7	15.8	79.0	4057	4			US-08-894-997-49	Sequence 49, App
C	8	15.8	79.0	19011	1		US-08-310-356-36	Sequence 36, App
C	9	15.8	79.0	19557	5		PCT-US92-06300-1	Sequence 1, Appl
C	10	15.2	76.0	201	4		US-09-134-001C-2584	Sequence 2584, App
C	11	15.2	76.0	313	4		US-09-221-017B-84	Sequence 84, App
C	12	15.2	76.0	685	4		US-09-122-400B-10	Sequence 10, App
C	13	15.2	76.0	703	4		US-09-221-017B-598	Sequence 598, App
C	14	15.2	76.0	2181	2		US-08-737-371A-1	Sequence 1, Appl
C	15	15.2	76.0	2181	5		PCT-US95-05853-1	Sequence 1, Appl
16	15.2	76.0	2251	4			US-08-991-677-11	Sequence 11, App
17	15.2	76.0	3026	4			US-09-149-476-317	Sequence 317, App
18	15.2	76.0	3054	4			US-09-149-476-194	Sequence 194, App
C	19	15.2	76.0	6496	4		US-09-221-017B-543	Sequence 543, App
20	15.2	76.0	18073	4			US-09-078-294-12	Sequence 12, App
21	15.2	76.0	59065	4			US-09-813-817-3	Sequence 3, Appl
22	15.2	76.0	59065	4			US-09-978-197-3	Sequence 3, Appl
23	15.2	76.0	246240	2			US-08-724-394A-20	Sequence 20, App
24	15.2	76.0	246240	2			US-08-724-394A-21	Sequence 21, App
25	15.2	76.0	246240	2			US-08-724-394A-22	Sequence 22, App
C	26	14.8	74.0	134	1		US-08-710-082-12	Sequence 12, App
C	27	14.8	74.0	134	3		US-08-913-462-12	Sequence 12, App

c	28	14.8	74.0	134	4	US-09-624-482-12	Sequence 12, Appl
c	29	14.8	74.0	154	3	US-08-710-082-13	Sequence 13, Appl
c	30	14.8	74.0	154	3	US-08-913-462-13	Sequence 13, Appl
c	31	14.8	74.0	154	4	US-09-624-482-13	Sequence 13, Appl
c	32	14.8	74.0	324	4	US-08-651-155B-22	Sequence 22, Appl
c	33	14.8	74.0	1180	4	US-09-276-531-51	Sequence 51, Appl
c	34	14.8	74.0	1260	2	US-08-747-381-1	Sequence 1, Appli
c	35	14.8	74.0	1260	3	US-08-969-761-1	Sequence 1, Appli
c	36	14.8	74.0	1260	3	US-09-225-443-1	Sequence 1, Appli
c	37	14.8	74.0	4973	4	US-09-381-862-6	Sequence 6, Appli
c	38	14.8	74.0	7278	4	US-09-091-219-1	Sequence 1, Appli
c	39	14.8	74.0	8657	4	US-08-961-527-44	Sequence 44, Appl
c	40	14.8	74.0	11056	4	US-09-004-838-23	Sequence 23, Appl
c	41	14.8	74.0	15062	4	US-09-004-838-89	Sequence 89, Appl
c	42	14.8	74.0	111282	4	US-09-754-250-3	Sequence 3, Appli
c	43	14.8	74.0	111282	4	US-09-754-250-3	Sequence 3, Appli
c	44	14.4	72.0	38	3	US-08-535-057A-9	Sequence 9, Appli
c	45	14.4	72.0	564	4	US-09-134-001C-2343	Sequence 2343, Ap

#### ALIGNMENTS

RESULT 1  
US-08-770-544-3  
; Sequence 3, Application US/08770544  
; Patent No. 5907085  
; GENERAL INFORMATION:  
; APPLICANT: Gonsalves, Dennis  
; APPLICANT: Ling, Kai-Siu  
; TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS PROTEINS AND  
; TITLE OF INVENTION: THEIR USES  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Nixon, Hargrave, Devans & Doyle LLP  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/770,544  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60009008  
; FILING DATE: 21-DEC-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldman, Michael L.  
; REGISTRATION NUMBER: 30,727  
; REFERENCE/DOCKET NUMBER: 19603/621  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1304  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1602 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-08-770-544-3  
Query Match 84.0%; Score 16.8; DB 2; Length 1602;  
Best Local Similarity 90.0%; Pred. No. 65;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 TTAGTCTTAGTTTAGTT 20

Db 1190 TTGCTCTAGTTTACT 1209

## RESULT 2

US-09-605-785-442/C  
Sequence 442, Application US/09605785  
Patent No. 6321716  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yugu  
APPLICANT: Henderson, Robert A.  
APPLICANT: Kalos, Michael D.  
APPLICANT: Ronger, Gary R.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darrick  
APPLICANT: Li, Samuel  
APPLICANT: Wang, Aijun  
APPLICANT: Skelky, Yasir A.W.  
APPLICANT: Hepler, William  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.427C16  
CURRENT APPLICATION NUMBER: US/09/605,785  
CURRENT FILING DATE: 2000-06-27  
NUMBER OF SEQ ID NOS: 835  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 442  
LENGTH: 362  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-605-785-442

Query Match 79.0%; Score 15.8; DB 4; Length 362;  
Best Local Similarity 89.5%; Pred. No. 1.7e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 TTAGTCTTACTTTTACT 19

Db 346 TTAATCTTAGTTTATT 328

## RESULT 3

US-09-439-313-442/C  
Sequence 442, Application US/09439313  
Patent No. 6329505  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan Louise  
APPLICANT: Jiang Yugu  
APPLICANT: Reed, Steven G.  
APPLICANT: Kalos, Michael  
APPLICANT: Ronger, Gary  
APPLICANT: Retter, Marc  
APPLICANT: Stolk, John  
APPLICANT: Day, Craig  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
FILE REFERENCE: 210121.427C9  
CURRENT APPLICATION NUMBER: US/09/439,313  
CURRENT FILING DATE: 1999-11-12  
NUMBER OF SEQ ID NOS: 575  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 442  
LENGTH: 362

TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-439-313-442

Query Match 79.0%; Score 15.8; DB 4; Length 362;  
Best Local Similarity 89.5%; Pred. No. 1.7e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 TTAGTCTTACTTTTACT 19

Db 346 TTAATCTTAGTTTATT 328

## RESULT 4

US-09-352-616A-442/C  
Sequence 442, Application US/09352616A  
Patent No. 6395278  
GENERAL INFORMATION:  
APPLICANT: Dillon, Davin C.  
APPLICANT: Harlocker, Susan Louise  
APPLICANT: Jiang, Yugu  
APPLICANT: Xu, Jiangchun  
APPLICANT: Mitcham, Jennifer Lynn  
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS  
FILE REFERENCE: 210121.427C8  
CURRENT APPLICATION NUMBER: US/09/352,616A  
CURRENT FILING DATE: 1999-07-13  
NUMBER OF SEQ ID NOS: 472  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 442  
LENGTH: 362  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-352-616A-442

Query Match 79.0%; Score 15.8; DB 4; Length 362;  
Best Local Similarity 89.5%; Pred. No. 1.7e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 TTAGTCTTACTTTTACT 19

Db 346 TTAATCTTAGTTTATT 328

## RESULT 5

US-09-556-877-46  
Sequence 46, Application US/09556877  
Patent No. 6432916  
GENERAL INFORMATION:  
APPLICANT: Probst, Peter  
APPLICANT: Bhatia, Ajay  
APPLICANT: Skelky, Yasir  
APPLICANT: Fling, Steve  
APPLICANT: Maisonneuve, Jeff  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
FILE REFERENCE: 210121.469C5  
CURRENT APPLICATION NUMBER: US/09/556,877  
CURRENT FILING DATE: 2000-04-19  
NUMBER OF SEQ ID NOS: 305  
SOFTWARE: FastSeq for Windows Version 3.0/4.0  
SEQ ID NO 46  
LENGTH: 427  
TYPE: DNA  
ORGANISM: Chlamydia  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (20)  
OTHER INFORMATION: n-A,T,C or G  
US-09-556-877-46

Query Match 79.0%; Score 15.8; DB 4; Length 427;

Best Local Similarity 89.5%; Pred. No. 1.7e+02;  
Matches 17: Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTTTACT 19  
|| |||||  
DB 30 TTTTCTTAGTTTACT 48

RESULT 6  
US-09-620-412C-46

; Sequence 46, Application US/09620412C  
; Patent No. 6448234  
; GENERAL INFORMATION:  
; APPLICANT: Steven P. Fling  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
; FILE REFERENCE: 210121.469C7  
; CURRENT APPLICATION NUMBER: US/09/620,412C  
; CURRENT FILING DATE: 2000-07-20  
; NUMBER OF SEQ ID NOS: 363  
; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 46  
; LENGTH: 427  
; TYPE: DNA  
; ORGANISM: Chlamydia  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (20)  
; OTHER INFORMATION: n-A,T,C or G  
US-09-620-412C-46

Query Match 79.0%; Score 15.8; DB 4; Length 427;  
Best Local Similarity 89.5%; Pred. No. 1.7e+02;  
Matches 17: Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTTTACT 19  
|| |||||  
DB 30 TTTTCTTAGTTTACT 48

RESULT 7  
US-08-894-997-49

; Sequence 49, Application US/08894997A  
; Patent No. 6270990  
; GENERAL INFORMATION:  
; APPLICANT: Schenbert, David J  
; TITLE OF INVENTION: NEURON-RESTRICTIVE SILENCER FACTOR  
; FILE REFERENCE: 17810-502 NRSE  
; CURRENT APPLICATION NUMBER: US/08/894,997A  
; CURRENT FILING DATE: 1998-01-06  
; EARLIER APPLICATION NUMBER: PCT/US96/02817  
; EARLIER FILING DATE: 1996-03-01  
; EARLIER APPLICATION NUMBER: 08/398,590  
; EARLIER FILING DATE: 1995-03-03  
; NUMBER OF SEQ ID NOS: 55  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 49  
; LENGTH: 4057  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: gene  
; LOCATION: (1)-(4057)  
; OTHER INFORMATION: Human NSRF  
US-08-894-997-49

Query Match 79.0%; Score 15.8; DB 4; Length 4057;  
Best Local Similarity 89.5%; Pred. No. 1.6e+02;  
Matches 17: Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 TTAGTCTTAGTTTACT 20  
|| ||||| || |||||

DB 3581 TAGTCTTAATTTTACT 3599

RESULT 8

US-08-310-356-36/C  
; Sequence 36, Application US/08310356  
; Patent No. 5648243  
; GENERAL INFORMATION:  
; APPLICANT: Hurwitz, David R  
; APPLICANT: Nathan, Margaret  
; APPLICANT: Shanl, Moshe  
; TITLE OF INVENTION: Transgenic Protein Production  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Legal Department  
; STREET: 500 Arcola Road  
; CITY: Collegeville  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19426  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Macintosh  
; OPERATING SYSTEM: Macintosh System 7.0  
; SOFTWARE: Microsoft Word Version 5.0 (Patentin)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/310,356  
; FILING DATE:  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/737,853  
; FILING DATE: 31-JUL-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goodman, Rosanne  
; REGISTRATION NUMBER: 32,534  
; REFERENCE/DOCKET NUMBER: A0856  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 454-3817  
; TELEFAX: (215) 454-3808  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19011 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 1737..1854  
; OTHER INFORMATION: /number= 1  
; OTHER INFORMATION: /citation= (11)  
; FEATURE:  
; NAME/KEY: intron  
; LOCATION: 1855..2563  
; OTHER INFORMATION: /number= 1  
; OTHER INFORMATION: /label= intron1  
; OTHER INFORMATION: /citation= (11)  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 2564..2621  
; OTHER INFORMATION: /number= 2  
; OTHER INFORMATION: /citation= (11)  
; FEATURE:  
; NAME/KEY: intron  
; LOCATION: 2622..4075  
; OTHER INFORMATION: /number= 2  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 4076..4208  
; OTHER INFORMATION: /number= 3  
; OTHER INFORMATION: /citation= (11)  
; FEATURE:  
; NAME/KEY: intron

LOCATION: 4209..6040  
OTHER INFORMATION: /number= 3  
FEATURE:  
NAME/KEY: exon  
LOCATION: 6041..6252  
OTHER INFORMATION: /number= 4  
FEATURE:  
NAME/KEY: intron  
LOCATION: 6253..6801  
OTHER INFORMATION: /number= 4  
FEATURE:  
NAME/KEY: exon  
LOCATION: 6802..6934  
OTHER INFORMATION: /number= 5  
FEATURE:  
NAME/KEY: intron  
LOCATION: 6935..7758  
OTHER INFORMATION: /number= 5  
FEATURE:  
NAME/KEY: exon  
LOCATION: 7759..7856  
OTHER INFORMATION: /number= 6  
FEATURE:  
NAME/KEY: intron  
LOCATION: 7857..9443  
OTHER INFORMATION: /number= 6  
FEATURE:  
NAME/KEY: exon  
LOCATION: 9444..9573  
OTHER INFORMATION: /number= 7  
FEATURE:  
NAME/KEY: intron  
LOCATION: 9574..10866  
OTHER INFORMATION: /number= 7  
FEATURE:  
NAME/KEY: exon  
LOCATION: 10867..11081  
OTHER INFORMATION: /number= 8  
FEATURE:  
NAME/KEY: intron  
LOCATION: 11082..12480  
OTHER INFORMATION: /number= 8  
FEATURE:  
NAME/KEY: exon  
LOCATION: 12481..12613  
OTHER INFORMATION: /number= 9  
FEATURE:  
NAME/KEY: intron  
LOCATION: 12614..13701  
OTHER INFORMATION: /number= 9  
FEATURE:  
NAME/KEY: exon  
LOCATION: 13702..13799  
OTHER INFORMATION: /number= 10  
FEATURE:  
NAME/KEY: intron  
LOCATION: 13800..14976  
OTHER INFORMATION: /number= 10  
FEATURE:  
NAME/KEY: exon  
LOCATION: 14977..15115  
OTHER INFORMATION: /number= 11  
FEATURE:  
NAME/KEY: intron  
LOCATION: 15116..15533  
OTHER INFORMATION: /number= 11  
FEATURE:  
NAME/KEY: exon  
LOCATION: 15534..15757  
OTHER INFORMATION: /number= 12  
FEATURE:  
NAME/KEY: intron  
LOCATION: 15758..16949

OTHER INFORMATION: /number= 12  
FEATURE:  
NAME/KEY: exon  
LOCATION: 16950..17082  
OTHER INFORMATION: /number= 13  
FEATURE:  
NAME/KEY: intron  
LOCATION: 17083..17696  
OTHER INFORMATION: /number= 13  
FEATURE:  
NAME/KEY: exon  
LOCATION: 17697..17764  
OTHER INFORMATION: /number= 14  
FEATURE:  
NAME/KEY: intron  
LOCATION: 17765..18534  
OTHER INFORMATION: /number= 14  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(1776..1854, 2564..2621, 4076..4208,  
LOCATION: 6041..6252, 6802..6934, 7759..7856, 9444..9573,  
LOCATION: 10867..11081, 12481..12613, 13702..13799,  
LOCATION: 14977..15115, 15534..15757, 16950..17082,  
LOCATION: 17697..17741)  
OTHER INFORMATION: /product= "human serum albumin"  
OTHER INFORMATION: /citation= (11)  
FEATURE:  
NAME/KEY: exon  
LOCATION: 18535..18697  
OTHER INFORMATION: /number= 15  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 17742..18697  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1737..1775  
PUBLICATION INFORMATION:  
AUTHORS: Minghetti, P P  
AUTHORS: Ruffner, D E  
AUTHORS: Kuang, W-J  
AUTHORS: Dennison, O E  
AUTHORS: Hawkins, J W  
AUTHORS: Beattie, W G  
AUTHORS: Dugaiczky, A  
TITLE: MOLECULAR STRUCTURE OF THE HUMAN ALBUMIN  
TITLE: GENE IS REVEALED BY NUCLEOTIDE SEQUENCE WITHIN  
JOURNAL: J. Biol. Chem.  
VOLUME: 261  
PAGES: 6747-6757  
DATE: 1986  
RELEVANT RESIDUES IN SEQ ID NO: 36: FROM 1 TO 19011  
US-08-310-356-36  
Query Match 79.0%; Score 15.8; DB 1; Length 19011;  
Best Local Similarity 89.5%; Pred. No. 1.5e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
OY 1 TTAGTCTTAGTTTACT 19  
||||||| ||| |||||  
Db 1409 TTAGTCTCAGCTTTAGT 1391  
RESULT 9  
PCT-US92-06300-1/c  
; Sequence 1, Application PC/TUS9206300  
; GENERAL INFORMATION:  
; APPLICANT: Hurwitz, David R  
; APPLICANT: Nathan, Margaret  
; APPLICANT: Shanli, Moshe  
; TITLE OF INVENTION: Transgenic Protein Production  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Rhone-Poulenc Rorer, Inc.  
STREET: 500 Virginia Ave., Bldg. 3A  
CITY: Ft. Washington  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19034  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/06300  
FILING DATE: 19920730  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Goodman, Rosanne  
REGISTRATION NUMBER: 52,534  
REFERENCE/DOCKET NUMBER: A0856-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 962-4130  
TELEFAX: (215) 962-4107  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19557 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
PUBLICATION INFORMATION:  
AUTHORS: Mindhetti, P P  
AUTHORS: Ruffner, D E  
AUTHORS: Kuang, W.-J.  
AUTHORS: Demnlson, O E  
AUTHORS: Hawkins, J W  
AUTHORS: Beattie, W G  
AUTHORS: Dugalczyk, A  
TITLE: MOLECULAR STRUCTURE OF THE HUMAN ALBUMIN GENE  
TITLE: IS REVEALED BY NUCLEOTIDE SEQUENCE WITHIN Q11-22  
JOURNAL: J. Biol. Chem.  
VOLUME: 261  
PAGES: 6747-6757  
DATE: 1986  
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 19002  
PCT-US92-06300-1  
Query Match 79.0%; Score 15.8; DB 5; Length 19557;  
Best Local Similarity 89.5%; Pred. No. 1.5e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
OY 1 TTAGTCTTACTTTTACT 19  
||||| ||| |||||  
Db 1964 TTAGTCTCAGTGTACT 1946  
RESULT 10  
US-09-134-001C-2584/c  
Sequence 2584, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14

NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 2584  
LENGTH: 201  
TYPE: DNA  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-2584  
Query Match 76.0%; Score 15.2; DB 4; Length 201;  
Best Local Similarity 85.0%; Pred. No. 3e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
OY 1 TTAGTCTTACTTTTACT 20  
||||| ||| |||||  
Db 115 TTGCTTCTCTTTTACT 96  
RESULT 11  
US-09-221-017B-84/c  
Sequence 84, Application US/09221017B  
Patent No. 644799  
GENERAL INFORMATION:  
APPLICANT: Ross, Bruce C.  
TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF  
NUMBER OF SEQUENCES: 1120  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & ROERSTER  
STREET: 755 PAGE MILL ROAD  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: FASTSEQ for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/221,017B  
FILING DATE: 23-DEC-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PP1182  
FILING DATE: 31-DEC-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PP1546  
FILING DATE: 30-JAN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PP2911  
FILING DATE: 09-APR-1998  
APPLICATION NUMBER: PCT/AU98/01023  
FILING DATE: 10-DEC-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Montoy, Gladys H  
REGISTRATION NUMBER: 32,430  
REFERENCE/DOCKET NUMBER: 27340-20021.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-813-5600  
TELEFAX: 650-494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 84:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 313 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: UNKNOWN  
ORIGINAL SOURCE:  
ORGANISM: PORPHYROMONAS GINGIVALIS  
FEATURE:  
NAME/KEY: misc\_feature

LOCATION: 1...313  
US-09-221-017B-84  
Query Match 76.0%; Score 15.2; DB 4; Length 313;  
Best Local Similarity 85.0%; Pred. No. 3e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 TTAGTCTCTAGTTTATTGTT 20  
DB 258 TTATTTCTGTGTTTGTGTT 239  
RESULT 12  
US-09-122-400B-10/c  
Sequence 10, Application US/09122400B  
Patent No. 6245974  
GENERAL INFORMATION:  
APPLICANT: Michalowski, Susan  
APPLICANT: Spiker, Steven  
TITLE OF INVENTION: MATRIX ATTACHMENT REGIONS  
FILE REFERENCE: Michalowski and Spiker  
CURRENT APPLICATION NUMBER: US/09/122,400B  
CURRENT FILING DATE: 1998-07-24  
PRIOR APPLICATION NUMBER: 60/066,118  
PRIOR FILING DATE: 1997-08-06  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 10  
LENGTH: 685  
TYPE: DNA  
ORGANISM: Nicotiana tabacum  
US-09-122-400B-10  
Query Match 76.0%; Score 15.2; DB 4; Length 685;  
Best Local Similarity 85.0%; Pred. No. 2.9e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 TTAGTCTCTAGTTTATTGTT 20  
DB 636 TTAGTTTCTGTTTATTGTTT 617  
RESULT 13  
US-09-221-017B-598/c  
Sequence 598, Application US/09221017B  
Patent No. 6444799  
GENERAL INFORMATION:  
APPLICANT: Ross, Bruce C.  
TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF  
NUMBER OF SEQUENCES: 1120  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 PAGE MILL ROAD  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: FastSeq for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/221,017B  
FILING DATE: 23-DEC-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PP1182  
FILING DATE: 31-DEC-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PP1546  
FILING DATE: 30-JAN-1998  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: PP2911  
FILING DATE: 09-APR-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU98/01023  
FILING DATE: 10-DEC-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Monroy, Gladys H  
REGISTRATION NUMBER: 32,430  
REFERENCE/DOCKET NUMBER: 27340-20021.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-813-5600  
TELEFAX: 650-494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 598:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 703 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: UNKNOWN  
ORIGINAL SOURCE:  
ORGANISM: PORPHYROMONAS GINGIVALIS  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1...703  
US-09-221-017B-598  
Query Match 76.0%; Score 15.2; DB 4; Length 703;  
Best Local Similarity 85.0%; Pred. No. 2.9e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 TTAGTCTCTAGTTTATTGTT 20  
DB 200 TTATTTCTGTTTGTGTTTGT 181  
RESULT 14  
US-08-737-371A-1/c  
Sequence 1, Application US/08737371A  
Patent No. 5959094  
GENERAL INFORMATION:  
APPLICANT: David WALLACH  
APPLICANT: Peter KUHNERT  
APPLICANT: Gotz EHRHARDT  
APPLICANT: Oliver KEMPER  
TITLE OF INVENTION: p75 TNF RECEPTOR PROMOTERS  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/737,371A  
FILING DATE: 08-NOVEMBER-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/05853  
FILING DATE: 11-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: IL 109,633  
FILING DATE: 11-MAY-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BROWDY, Roger L.  
REGISTRATION NUMBER: 25,618



REFERENCE/DOCKET NUMBER: WALLACH-14  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2181 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 2094..2171  
US-08-737-371A-1

Query Match 76.0%; Score 15.2; DB 2; length 2181;  
Best Local Similarity 85.0%; Pred. No. 2.8e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 TTAGTCTTAGTCTTACTT 20  
||| ||||| ||||| ||  
Db 1591 TTAATCTTATTTTAAAT 1572

## RESULT 15

PCT-US95-05853-1/c  
Sequence 1, Application PC/TUS9505853  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: p75 TNF RECEPTOR PROMOTERS  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/05853  
FILING DATE: 11-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: IL 109,633  
FILING DATE: 11-MAY-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: BROWDY, Roger L.  
REGISTRATION NUMBER: 25,618  
REFERENCE/DOCKET NUMBER: WALLACH-14 PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2181 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 2094..2171  
PCT-US95-05853-1

Query Match 76.0%; Score 15.2; DB 5; length 2181;  
Best Local Similarity 85.0%; Pred. No. 2.8e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 TTAGTCTTAGTCTTACTT 20  
||| ||||| ||||| ||  
Db 1591 TTAATCTTATTTTAAAT 1572

Search completed: April 6, 2003, 11:55:45  
Job time : 53.25 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 11:24:32 ; Search time 58 seconds  
(without alignments)  
302.472 Million cell updates/sec

Title: US-09-980-265-10

Perfect score: 20

Sequence: 1 ttactcttacttttactgt 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapept 1.0

Searched: 593429 seqs, 438583890 residues

Total number of hits satisfying chosen parameters: 1186858

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Database : Published\_Applications\_MA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*

2: /cgn2\_6/ptodata/1/pubpna/PC1\_NEW\_PUB.seq:\*

3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*

4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*

5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*

6: /cgn2\_6/ptodata/1/pubpna/PC1US\_PUBCOMB.seq:\*

7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*

8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*

9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*

10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*

11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*

12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*

13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*

14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16.8	84.0	548	10	US-09-924-035A-396
2	16.8	84.0	63000	10	US-09-780-172-18
3	16.4	82.0	50000	9	US-10-060-763-4
4	16.4	82.0	50000	12	US-10-063-763-4
5	16	80.0	219	10	US-09-998-588-2498
6	15.8	79.0	362	9	US-10-012-896-442
7	15.8	79.0	362	9	US-09-895-793-442
8	15.8	79.0	362	9	US-09-895-814-442
9	15.8	79.0	362	10	US-09-759-143-442
10	15.8	79.0	362	10	US-09-780-669-442
11	15.8	79.0	362	10	US-09-822-827-442
12	15.8	79.0	427	10	US-09-841-132-46
13	15.8	79.0	442	10	US-09-880-107-2657
14	15.8	79.0	579	10	US-09-841-132-546
15	15.8	79.0	24132	10	US-09-764-860-661
16	15.8	79.0	513509	9	US-09-754-853A-4
17	15.8	79.0	640681	10	US-09-790-988-1
18	15.4	77.0	269	10	US-09-294-093B-1039
19	15.4	77.0	1505	10	US-09-822-849A-138

20	15.4	77.0	2000	9	US-09-938-842A-4209	Sequence 4209, Ap
21	15.4	77.0	659158	9	US-09-771-208-20	Sequence 20, Appl
22	15.2	76.0	192	10	US-09-815-242-2912	Sequence 2912, Ap
23	15.2	76.0	192	10	US-09-815-242-2918	Sequence 2918, Ap
24	15.2	76.0	192	10	US-09-815-242-2947	Sequence 2947, Ap
25	15.2	76.0	192	10	US-09-815-242-3417	Sequence 3417, Ap
26	15.2	76.0	192	10	US-09-815-242-3488	Sequence 3488, Ap
27	15.2	76.0	256	10	US-09-783-590-9439	Sequence 9439, Ap
28	15.2	76.0	368	9	US-09-232-785-352	Sequence 352, App
29	15.2	76.0	378	10	US-09-815-343-1089	Sequence 1089, Ap
30	15.2	76.0	397	10	US-09-867-701-5873	Sequence 5873, Ap
31	15.2	76.0	424	9	US-09-796-692-4100	Sequence 4100, Ap
32	15.2	76.0	462	10	US-09-960-352-7660	Sequence 7660, Ap
33	15.2	76.0	494	10	US-09-867-701-6731	Sequence 6731, Ap
34	15.2	76.0	572	10	US-09-864-761-12035	Sequence 12035, A
35	15.2	76.0	620	9	US-09-908-931B-37	Sequence 37, Appl
36	15.2	76.0	667	9	US-09-908-931B-12	Sequence 12, Appl
37	15.2	76.0	685	10	US-09-816-894-10	Sequence 10, Appl
38	15.2	76.0	697	10	US-09-770-149-277	Sequence 277, App
39	15.2	76.0	701	10	US-09-770-149-251	Sequence 251, App
40	15.2	76.0	760	9	US-10-040-739-1332	Sequence 1332, Ap
41	15.2	76.0	1019	9	US-10-278-173-153	Sequence 153, App
42	15.2	76.0	1070	9	US-10-278-173-133	Sequence 133, App
43	15.2	76.0	1202	10	US-09-822-830A-591	Sequence 591, App
44	15.2	76.0	1557	9	US-10-001-873-1	Sequence 1, Appl
45	15.2	76.0	1707	9	US-10-108-605-152	Sequence 152, App

## ALIGNMENTS

RESULT 1  
US-09-924-035A-396/C  
Sequence 396, Application US/09924035A  
Patent No. US20020142319A1  
GENERAL INFORMATION:  
APPLICANT: Grilach, Jim  
TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
FILE REFERENCE: 2011US  
CURRENT FILING DATE: 2000-08-11  
PRIOR APPLICATION NUMBER: US 60/148,784  
PRIOR FILING DATE: 1999-08-13  
NUMBER OF SEQ ID NOS: 900  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 396  
LENGTH: 548  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
NAME/KEY: misc-feature  
LOCATION: (1)...(548)  
OTHER INFORMATION: n = A,T,C or G  
US-09-924-035A-396

Query Match 84.0%; Score 16.8; DB 10; Length 548;  
Best Local Similarity 90.0%; Pred. No. 3,4e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

DB 172 TTAGTCTTACTTGTAGTT 153  
172 TTAGTCTTACTTGTAGTT 20

RESULT 2  
US-09-780-172-18  
Sequence 18, Application US/09780172  
Patent No. US20020147163A1  
GENERAL INFORMATION:  
APPLICANT: Robert McKay  
APPLICANT: Susan M. Freiler  
APPLICANT: Jacqueline Wyatt

```

; TITLE OF INVENTION: ANTISENSE MODULATION OF CASEIN KINASE 2-ALPHA EXPRESSION
; FILE REFERENCE: RFS-0159
; CURRENT APPLICATION NUMBER: US/09/780,172
; CURRENT FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 96
; SEQ ID NO 18
; LENGTH: 63000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-780-172-18

Query Match      84.0%; Score 16.8; DB 10; Length 63000;
Best Local Similarity 90.0%; Pred. No. 5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TTACTCTTACTTTTACTT 20
    1 ||||| ||||| ||||| |||||
Db 18859 TTAGTCTTAACTTTTACTT 18878

RESULT 3
US-10-060-763-4
; Sequence 4, Application US/10060763
; Publication No. US20030022286A1
; GENERAL INFORMATION:
; APPLICANT: Curtlis, Rory A.J.
; TITLE OF INVENTION: No. US20030022286A1el Transporter-Like Genes and Uses Therefor
; FILE REFERENCE: 10147-8
; CURRENT APPLICATION NUMBER: US/10/060,763
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 4
; LENGTH: 50000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-060-763-4

Query Match      82.0%; Score 16.4; DB 9; Length 50000;
Best Local Similarity 94.4%; Pred. No. 7.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTACTCTTACTTTTACTT 18
    1 ||||| ||||| ||||| |||||
Db 4192 TTACTCTTACTTTTACTT 4209

RESULT 4
US-10-063-763-4
; Sequence 4, Application US/10063763
; Patent No. US2001000163A1
; GENERAL INFORMATION:
; APPLICANT: Curtlis, Rory A.J.
; TITLE OF INVENTION: No. US2001000163A1el Transporter-Like Genes and Uses Therefor
; FILE REFERENCE: 10147-8
; CURRENT APPLICATION NUMBER: US/10/063,763
; CURRENT FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: 09/365,162
; PRIOR FILING DATE: 1999-JUL-30
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 4
; LENGTH: 50000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-063-763-4

Query Match      82.0%; Score 16.4; DB 12; Length 50000;
Best Local Similarity 94.4%; Pred. No. 7.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTTTTACTT 18
```

```

Db 4192 TTACTCTTACTTTTACTT 4209
    1 ||||| ||||| ||||| |||||

RESULT 5
US-09-998-598-2498
; Sequence 2498, Application US/09998598
; Patent No. US20020150922A1
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Chenault, Ruth A.
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.561
; CURRENT APPLICATION NUMBER: US/09/998,598
; CURRENT FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 2606
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 2498
; LENGTH: 219
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-998-598-2498

Query Match      80.0%; Score 16; DB 10; Length 219;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TTCTTACTTTTACTT 20
    5 ||||| ||||| ||||| |||||
Db 143 TTCTTACTTTTACTT 158

RESULT 6
US-10-012-896-442/C
; Sequence 442, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqin
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedavick, Thomas S.
; APPLICANT: Carter, Derrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hurai, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Roy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Mantanabe, Yoshihiro
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012,896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 442
; LENGTH: 362
```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-896-442

```

```

Query Match          79.0%; Score 15.8; DB 9; Length 362;
Best Local Similarity 89.5%; Pred. No. 8.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 1 TTAGTCTTAGTTTACT 19
    ||| ||||| ||||| |||
Db 346 TTAATCTTAGTTTACT 328

```

```

RESULT 7
US-09-895-793-442/c
; Sequence 442, Application US/09895793
; Publication No. US20020192763A1
GENERAL INFORMATION:

```

```

APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Yajun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hurai, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Bassols, Carlota
APPLICANT: Foy, Teresa
APPLICANT: Fanger, Gary R.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.534C2
CURRENT APPLICATION NUMBER: US/09/895,793
CURRENT FILING DATE: 2001-06-29
NUMBER OF SEQ ID NOS: 982
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 442
LENGTH: 362
TYPE: DNA
ORGANISM: Homo sapiens
US-09-895-793-442

```

```

Query Match          79.0%; Score 15.8; DB 9; Length 362;
Best Local Similarity 89.5%; Pred. No. 8.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 1 TTAGTCTTAGTTTACT 19
    ||| ||||| ||||| |||
Db 346 TTAATCTTAGTTTACT 328

```

```

RESULT 8
US-09-895-814-442/c
; Sequence 442, Application US/09895814
; Publication No. US20020193296A1
GENERAL INFORMATION:

```

```

APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu
APPLICANT: Kalos, Michael D.

```

```

APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Yajun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hurai, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Bassols, Carlota
APPLICANT: Foy, Teresa
APPLICANT: Fanger, Gary R.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C26
CURRENT APPLICATION NUMBER: US/09/895,814
CURRENT FILING DATE: 2001-06-29
NUMBER OF SEQ ID NOS: 990
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 442
LENGTH: 362
TYPE: DNA
ORGANISM: Homo sapiens
US-09-895-814-442

```

```

Query Match          79.0%; Score 15.8; DB 9; Length 362;
Best Local Similarity 89.5%; Pred. No. 8.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 1 TTACTCTTAGTTTACT 19
    ||| ||||| ||||| |||
Db 346 TTAATCTTAGTTTACT 328

```

```

RESULT 9
US-09-759-143-442/c
; Sequence 442, Application US/09759143
; Patent No. US2002002248A1
GENERAL INFORMATION:

```

```

APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel
APPLICANT: Wang, Yajun
APPLICANT: Skeiky, Yasir A.W.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C23
CURRENT APPLICATION NUMBER: US/09/759,143
CURRENT FILING DATE: 2001-01-12
NUMBER OF SEQ ID NOS: 934
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 442
LENGTH: 362
TYPE: DNA
ORGANISM: Homo sapiens
US-09-759-143-442

```

Query Match 79.0%; Score 15.8; DB 10; Length 362;  
Best Local Similarity 89.5%; Pred. No. 8.3e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTTTTACT 19  
||| ||||| ||||| |||  
Db 346 TTAATCTTACTTTTATT 328

RESULT 10  
US-09-780-669-442/C  
; Sequence 442, Application US/09780669  
; Patent No. US20020051977A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Devin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqi  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darlick  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skelky, Yasir A.W.  
; APPLICANT: Hepler, William  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghon, Raymond L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.427C24  
; CURRENT APPLICATION NUMBER: US/09/780.669  
; CURRENT FILING DATE: 2001-02-09  
; NUMBER OF SEQ ID NOS: 943  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 442  
; LENGTH: 362  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-780-669-442

Query Match 79.0%; Score 15.8; DB 10; Length 362;  
Best Local Similarity 89.5%; Pred. No. 8.3e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTTTTACT 19  
||| ||||| ||||| |||  
Db 346 TTAATCTTACTTTTATT 328

RESULT 11  
US-09-822-827-442/C  
; Sequence 442, Application US/09822827  
; Patent No. US20020081680A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.534C1  
; CURRENT APPLICATION NUMBER: US/09/822.827  
; CURRENT FILING DATE: 2001-03-28  
; NUMBER OF SEQ ID NOS: 982  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 442  
; LENGTH: 362  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-09-822-827-442  
Query Match 79.0%; Score 15.8; DB 10; Length 362;  
Best Local Similarity 89.5%; Pred. No. 8.3e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTTTTACT 19  
||| ||||| ||||| |||  
Db 346 TTAATCTTACTTTTATT 328

RESULT 12  
US-09-841-132-46  
; Sequence 46, Application US/09841132  
; Patent No. US20020061848A1  
; GENERAL INFORMATION:  
; APPLICANT: Bhattacharya, Ajay  
; APPLICANT: Skelky, Yasir A.W.  
; APPLICANT: Probst, Peter  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
; FILE REFERENCE: 210121.469C8  
; CURRENT APPLICATION NUMBER: US/09/841.132  
; CURRENT FILING DATE: 2001-04-23  
; NUMBER OF SEQ ID NOS: 599  
; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 46  
; LENGTH: 427  
; TYPE: DNA  
; ORGANISM: Chlamydia  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (20)  
; OTHER INFORMATION: n=A,T,C or G  
US-09-841-132-46

Query Match 79.0%; Score 15.8; DB 10; Length 427;  
Best Local Similarity 89.5%; Pred. No. 8.4e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTTTTACT 19  
||| ||||| ||||| |||  
Db 30 TTTTCTTACTTTTACT 48

RESULT 13  
US-09-880-107-2657/C  
; Sequence 2657, Application US/09880107  
; Patent No. US20020142981A1  
; GENERAL INFORMATION:  
; APPLICANT: Horne, Darci T.  
; APPLICANT: Vockley, Joseph G.  
; APPLICANT: Scherf, Uwe  
; APPLICANT: Gene Logic, Inc.  
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
; FILE REFERENCE: 44921-5028-WO  
; CURRENT APPLICATION NUMBER: US/09/880.107  
; CURRENT FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/211,379  
; PRIOR FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: US 60/237,054  
; PRIOR FILING DATE: 2000-10-02  
; NUMBER OF SEQ ID NOS: 3950  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2657  
; LENGTH: 442  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 N70481  
US-09-880-107-2657

Query Match 79.0%; Score 15.8; DB 10; Length 442;

Best Local Similarity 89.5%; Pred. No. 8.4e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Oy 2 TAGTTCCTAGTTTCTT 20  
 |||||||  
 Db 103 TAGTTCCTAATTATT 85

RESULT 14  
 US-09-841-132-546/C  
 ; Sequence 546, Application US/09841132  
 ; Patent No. US20020061848A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bhatia, Ajay  
 ; APPLICANT: Skeiky, Yasir A.W.  
 ; APPLICANT: Probst, Peter  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
 ; FILE REFERENCE: 210121.469C8  
 ; CURRENT APPLICATION NUMBER: US/09/841.132  
 ; CURRENT FILING DATE: 2001-04-23  
 ; NUMBER OF SEQ ID NOS: 599  
 ; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
 ; SEQ ID NO 546  
 ; LENGTH: 579  
 ; TYPE: DNA  
 ; ORGANISM: C. Trachomatis D serovar  
 US-09-841-132-546

Query Match 79.0%; Score 15.8; DB 10; Length 579;  
 Best Local Similarity 89.5%; Pred. No. 8.6e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 TTAGTTCCTAGTTTCTT 19  
 |||||||  
 Db 579 TTTTCTTCTAGTTTCTT 561

RESULT 15  
 US-09-764-860-661  
 ; Sequence 661, Application US/09764860  
 ; Patent No. US20020094953A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: PC008  
 ; CURRENT APPLICATION NUMBER: US/09/764.860  
 ; CURRENT FILING DATE: 2001-01-17  
 ; Prior application data removed - consult PALM or file wrapper  
 ; NUMBER OF SEQ ID NOS: 1198  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 661  
 ; LENGTH: 24132  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-764-860-661

Query Match 79.0%; Score 15.8; DB 10; Length 24132;  
 Best Local Similarity 89.5%; Pred. No. 1.2e+03;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 2 TAGTTCCTAGTTTCTT 20  
 |||||||  
 Db 13693 TAGTTCCTAATTATT 13711

Search completed: April 6, 2003, 12:09:57  
 Job time : 86 secs

**THIS PAGE BLANK (USPTO)**





JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Unpublished  
2 (bases 1 to 155877)  
Morley, K.C.  
Direct Submission  
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GYWL  
Center clone name: CH230-117D11  
----- Summary Statistics  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 73163 bases at least Q40  
Consensus quality: 82138 bases at least Q30  
Consensus quality: 87186 bases at least Q20

-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 85 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1  
1063 1162: contig of 1062 bp in length  
1163 1162: gap of unknown length  
2172 2271: contig of 1009 bp in length  
2271 2272: gap of unknown length  
2272 3321: contig of 1050 bp in length  
3321 3421: gap of unknown length  
3421 4463: contig of 1042 bp in length  
4463 4563: gap of unknown length  
4563 5587: contig of 1024 bp in length  
5587 5687: gap of unknown length  
5687 6712: contig of 1025 bp in length  
6712 6812: gap of unknown length  
6812 7823: contig of 1010 bp in length  
7823 7923: gap of unknown length  
7923 9753: contig of 1831 bp in length  
9753 9854: gap of unknown length  
9854 10888: contig of 1034 bp in length  
10888 10987: gap of unknown length  
10987 12401: contig of 1414 bp in length  
12401 12501: gap of unknown length  
12501 14037: contig of 1536 bp in length  
14037 14137: gap of unknown length  
14137 15197: contig of 1060 bp in length  
15197 15297: gap of unknown length  
15297 16644: contig of 1347 bp in length  
16644 16744: gap of unknown length  
16744 18177: contig of 1433 bp in length  
18177 18278: gap of unknown length  
18278 19658: contig of 1381 bp in length  
19658 19758: gap of unknown length  
19758 20806: contig of 1048 bp in length  
20806 20906: gap of unknown length  
20906 22312: contig of 1406 bp in length  
22312 22412: gap of unknown length  
22412 23922: contig of 1510 bp in length  
23922 24022: gap of unknown length  
24022 25433: contig of 1411 bp in length  
25433 25533: gap of unknown length  
25533 26809: contig of 1276 bp in length  
26809 26909: gap of unknown length

26910 28404: contig of 1495 bp in length  
28404 28405: gap of unknown length  
28405 28505: contig of 1166 bp in length  
28505 29671: gap of unknown length  
29671 29771: contig of 1224 bp in length  
29771 30995: gap of unknown length  
30995 31095: contig of 1581 bp in length  
31095 32675: gap of unknown length  
32675 34131: contig of 1355 bp in length  
34131 34230: gap of unknown length  
34230 35489: contig of 1259 bp in length  
35489 35589: gap of unknown length  
35589 37026: contig of 1437 bp in length  
37026 37126: gap of unknown length  
37126 38797: contig of 1670 bp in length  
38797 38896: gap of unknown length  
38896 40001: contig of 1105 bp in length  
40001 40002: gap of unknown length  
40002 40102: contig of 1014 bp in length  
40102 41116: gap of unknown length  
41116 41216: contig of 1354 bp in length  
41216 42570: gap of unknown length  
42570 43831: contig of 1161 bp in length  
43831 43930: gap of unknown length  
43930 45281: contig of 1351 bp in length  
45281 45381: gap of unknown length  
45381 47441: contig of 2060 bp in length  
47441 47541: gap of unknown length  
47541 49026: contig of 1485 bp in length  
49026 49126: gap of unknown length  
49126 50263: contig of 1137 bp in length  
50263 50363: gap of unknown length  
50363 52441: contig of 2078 bp in length  
52441 52541: gap of unknown length  
52541 53981: contig of 1440 bp in length  
53981 54081: gap of unknown length  
54081 55502: contig of 1421 bp in length  
55502 55602: gap of unknown length  
55602 55893: contig of 1291 bp in length  
55893 56993: gap of unknown length  
56993 58954: contig of 1960 bp in length  
58954 59054: gap of unknown length  
59054 60351: contig of 1298 bp in length  
60351 60451: gap of unknown length  
60451 61848: contig of 1397 bp in length  
61848 61948: gap of unknown length  
61948 64132: contig of 2184 bp in length  
64132 64232: gap of unknown length  
64232 65624: contig of 1392 bp in length  
65624 65724: gap of unknown length  
65724 66793: contig of 1071 bp in length  
66793 66893: gap of unknown length  
66893 68161: contig of 1266 bp in length  
68161 68261: gap of unknown length  
68261 69390: contig of 1129 bp in length  
69390 69490: gap of unknown length  
69490 70785: contig of 1295 bp in length  
70785 70885: gap of unknown length  
70885 73269: contig of 2384 bp in length  
73269 73369: gap of unknown length  
73369 75050: contig of 1681 bp in length  
75050 75150: gap of unknown length  
75150 76285: contig of 1135 bp in length  
76285 76385: gap of unknown length  
76385 78166: contig of 1781 bp in length  
78166 78266: gap of unknown length  
78266 80185: contig of 1919 bp in length  
80185 80285: gap of unknown length  
80285 81681: contig of 1396 bp in length  
81681 81781: gap of unknown length  
81781 83340: contig of 1559 bp in length  
83340 83440: gap of unknown length  
83440 85436: contig of 1996 bp in length

```

* 85437 85536: gap of unknown length

Query Match      100.0%; Score 20; DB 1; Length 15587;
Best Local Similarity 100.0%; Pred. No. 2,2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTGTTCTTACT 20
    |||||
Db 46393 TTAGTCTTACTGTTCTTACT 46412

RESULT 2
AC015120
LOCUS      AC015120
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, In ordered
           pieces.
ACCESSION  AC015120
VERSION    AC015120.1 GI:6436215
KEYWORDS   HTG; HTGS_PHASE2.
SOURCE     Drosophila melanogaster.
ORGANISM   Drosophila melanogaster.
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1 (bases 1 to 19997)
AUTHORS   Adams, M. and Venter, J.C.
TITLE     Direct Submission
JOURNAL   Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,
            Rockville, MD, USA
COMMENT   This sequence was identified as CDM:10211942 by the submitter.
            For further information on this sequence e-mail to fly@celera.com.
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.

FEATURES
    source
        1..19997
           /organism="Drosophila melanogaster"
           /db_xref="taxon:7227"

BASE COUNT  5276 a 4312 c 4203 g 6206 t

ORIGIN
Query Match      95.0%; Score 19; DB 1; Length 19997;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTGTTCTTACT 19
    |||||
Db 1511 TTAGTCTTACTGTTCTTACT 1529

RESULT 3
AC010951/c
LOCUS      AC010951
DEFINITION Homo sapiens clone RP11-3N19, WORKING DRAFT SEQUENCE, 8 unordered
           pieces.
ACCESSION  AC010951
VERSION    AC010951.4 GI:7341835
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE     Homo sapiens.
ORGANISM   Homo sapiens.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 148208)
AUTHORS   Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE     Homo sapiens, clone RP11-3N19
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 148208)
AUTHORS   Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
            Baldwin, J., Barna, N., Becker, R., Boguslavsky, L., Bokhale, B.,
            Brown, A., Castle, A., Colangelo, M., Collins, S., Collumore, A.,
            Cooke, P., Dearfiano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
            Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,

```

```

Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howard, J.C., Johnson, R., Jones, C., Kann, L., Karats, A., Klein, J.,
Lehoczky, J., Lien, C., Locke, K., Macdonald, P., Marquis, N.,
McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meltrin, J.,
Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tessfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, D., Wu, X.,
Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.

Direct Submission
Submitted (28-SEP-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 30, 2000 this sequence version replaced gi:6087956.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center -----
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information -----
Center project name: L2785
Center clone name: 3_N_19

----- Summary Statistics -----
Sequencing vector: M13; M7815: 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 142002 bases at least Q40
Consensus quality: 144771 bases at least Q30
Consensus quality: 145833 bases at least Q20
Insert size: 147000; agarose-fp
Insert size: 147508; sum-of-contigs
Quality coverage: 4.9 in Q20 bases; sum-of-contigs
Quality coverage: 4.9 in Q20 bases; sum-of-contigs

-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2007: contig of 2007 bp in length
* 2008 2107: gap of 100 bp
* 2108 4935: contig of 2828 bp in length
* 4936 5035: gap of 100 bp
* 5036 18397: contig of 13362 bp in length
* 18398 18497: gap of 100 bp
* 18498 33384: contig of 14887 bp in length
* 33385 33484: gap of 100 bp
* 33485 46070: contig of 12586 bp in length
* 46071 46170: gap of 100 bp
* 46171 66736: contig of 20566 bp in length
* 66737 66836: gap of 100 bp
* 66837 105893: contig of 39057 bp in length
* 105894 105993: gap of 100 bp
* 105994 148208: contig of 42215 bp in length.

FEATURES
    source
        1..148208
           /organism="Homo sapiens"
           /db_xref="taxon:9606"
           /clone="RP11-3N19"
           /clone_11b="RP11-11 Human Male BAC"
           1..2007
              /note="assembly_fragment"
              2108..4935
                 /note="assembly_fragment"
                 5036..18397
                    /note="assembly_fragment"
                    clone_end:7
                    vector_side:right"

```



```

source
1. 173471
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP23-251G19"
/clone_lib="RPC1-23 Female Mouse BAC"
1. 1586
/misc_feature
/note="assembly-fragment"
1687. 43654
/misc_feature
/note="assembly-fragment"
4375. 47502
/misc_feature
/note="assembly-fragment"
47603. 51368
/misc_feature
/note="assembly-fragment"
51469. 57076
/misc_feature
/note="assembly-fragment"
57177. 61964
/misc_feature
/note="assembly-fragment"
62065. 73715
/misc_feature
/note="assembly-fragment"
73816. 87823
/misc_feature
/note="assembly-fragment"
87924. 99375
/misc_feature
/note="assembly-fragment"
99476. 121365
/misc_feature
/note="assembly-fragment"
121466. 146521
/misc_feature
/note="assembly-fragment"
146622. 172750
/misc_feature
/note="assembly-fragment"
172851. 173471
/misc_feature
/note="assembly-fragment"
clone_end:77
vector_side:right"

BASE COUNT 55827 a 33466 c 33802 g 49176 t 1200 others
ORIGIN

Query Match 95.0%; Score 19; DB 1; Length 173471;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TAGTCTTAGTCTTAGT 20
|||||
Db 112559 TAGTCTTAGTCTTAGT 112577

RESULT 5
AC130415 186581 bp DNA linear HTG 10-AUG-2002
DEFINITION Homo sapiens chromosome 12 clone RP11-90C1, *** SEQUENCING IN
ACCESSION AC130415
VERSION AC130415.1 GI:22203172
KEYWORDS HTG; HTGS_PHASE1.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Muzny D.M., Adams C., Adio-Oduola B., Ali-osman, F.R., Allen C.,
Alibrooks S.L., Amaralunge H.C., Are J.R., Ayele M., Banks T.,
Barbata J., Benton J., Blinag K., Blankenburg K., Bonin D.,
Bouck J., Bowle S., Brileva M., Brown E., Brown M., Bryant N.P.,
Buhay C., Burch P., Burkett C., Burrell K.L., Byrd N.C.,
Carroll T.F., Carter M., Cavazos S.R., Chacko J., Chavez D.,
Chen G., Chen R., Chen Z., Chowdhry I., Christopoulos C.,
Cleveland C.D., Cox C., Coyle M.D., Dathorne S.R., David R.,
DeJana M.L., Davis C., Davy-Carroll L., Dederich D.A.,
DeJana K.R., Delgado O., Denn A.L., Ding Y., Dinh H.H.,
Douthwaite K.J., Draper H., Dugan-Rocha S., Durbin K.J.,
Eamhart C., Edgar D., Edwards C.C., Elhaj C., Escotto M.,
Falls T., Ferraguto D., Flagg N., Ford J., Foster P., Frantz P.,
Gabisi A., Gao J., Garcia A., Garner T., Garza N., Gill R.,
Gorrell J.H., Guevara W., Gunaratne P., Hale S., Hamilton K.,

```

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

COMMENT

```

Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hughes, M., Holloway, C., Hollins, B.,
Homi, F., Howard, S., Huber, J., Hulik, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karlssohn, E., Kelly, S., Khan, U., King, L., Kovach, J., Kovar, C.,
Kratovic, U., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Liu, C., Liu, J., Liu, W., Louised, R.,
Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Melzer, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenkwu, S., Ogih, M., Okwuonu, G.,
Oraguwe, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Scherer, S., Scott, G., Shen, H., Shooshitari, N., Sisson, T.,
Sodergren, E., Sonalike, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Umanli, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wlezyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 16581)
Morley, K.C.
Direct Submission
Submitted (10-AUG-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HCVU
Center clone name: RP11-90C1
----- Summary Statistics
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 177555 bases at least Q40
Consensus quality: 179936 bases at least Q30
Consensus quality: 181287 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 19 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1 2175: contig of 2175 bp in length
2176 2275: gap of unknown length
2276 4405: contig of 2130 bp in length
4405 4505: gap of unknown length
4506 8401: contig of 3896 bp in length
8402 8501: gap of unknown length
8502 12794: contig of 4293 bp in length
12795 12894: gap of unknown length
12895 15194: contig of 2300 bp in length
15195 15294: gap of unknown length
15295 20240: contig of 4946 bp in length
20241 20340: gap of unknown length
20341 24892: contig of 4552 bp in length
24893 24992: gap of unknown length
24993 30199: contig of 5207 bp in length

```

30200 30299: gap of unknown length  
30300 33759: contig of 3460 bp in length  
33760 33859: gap of unknown length  
33860 38788: contig of 4929 bp in length  
38789 38888: gap of unknown length  
38889 45969: contig of 7081 bp in length  
45970 46069: gap of unknown length  
46070 55234: contig of 9165 bp in length  
55235 55334: gap of unknown length  
55335 66750: contig of 11416 bp in length  
66751 81272: contig of 14422 bp in length  
81273 81372: gap of unknown length  
81373 95615: contig of 14243 bp in length  
95616 95715: gap of unknown length  
95716 109015: contig of 13300 bp in length  
109016 109115: gap of unknown length  
109116 125169: contig of 16054 bp in length  
125170 125269: gap of unknown length  
125270 146580: contig of 21311 bp in length  
146581 146680: gap of unknown length  
146681 186581: contig of 39901 bp in length.  
Location/Qualifiers  
1. 186581  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="RP11-90C1"

BASE COUNT 58039 a 33763 c 32807 g 60139 t 1833 others  
ORIGIN

Query Match 95.0%; Score 19; DB 1; Length 186581;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TTACTTCTAGTCTTACTT 20  
|||||  
Db 53280 TAGTCTTACTTCTTACTT 53298

RESULT 6  
AC019958/c  
LOCUS  
DEFINITION Drosophila melanogaster, \*\*\* SEQUENCING IN PROGRESS \*\*\*, in ordered pieces.  
AC019958  
VERSION AC019958.1 GI:6664939  
KEYWORDS HTG: HTGS\_PHASE2.  
SOURCE Drosophila melanogaster.  
ORGANISM Drosophila melanogaster.  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 5421)  
Adams, M. and Venter, J.C.  
Direct Submission  
Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA  
This sequence was identified as CDM:10211530 by the submitter.  
For more information on this record e-mail to fly@celera.com.  
\* NOTE: This is a 'working draft' sequence.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

FEATURES  
source  
1. 5421  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"

BASE COUNT 1558 a 1150 c 1146 g 1567 t  
ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 5421;  
Best Local Similarity 95.0%; Pred. No. 14;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 TTACTTCTAGTCTTACTT 20  
|||||  
Db 435 TTACTTCTAGTCTTACTT 416

RESULT 7  
AC015356/c  
LOCUS  
DEFINITION Drosophila melanogaster, \*\*\* SEQUENCING IN PROGRESS \*\*\*, in ordered pieces.  
AC015356  
VERSION AC015356.1 GI:6435979  
KEYWORDS HTG: HTGS\_PHASE2.  
SOURCE Drosophila melanogaster.  
ORGANISM Drosophila melanogaster.  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 14921)  
Adams, M. and Venter, J.C.  
Direct Submission  
Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA  
This sequence was identified as CDM:1021151 by the submitter.  
For further information on this sequence e-mail to fly@celera.com.  
\* NOTE: This is a 'working draft' sequence.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

FEATURES  
source  
1. 14921  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"

BASE COUNT 4439 a 3179 c 2922 g 4381 t  
ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 14921;  
Best Local Similarity 95.0%; Pred. No. 16;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTACTTCTAGTCTTACTT 20  
|||||  
Db 12619 TTACTTCTAGTCTTACTT 12600

RESULT 8  
AC019964  
LOCUS  
DEFINITION Drosophila melanogaster, \*\*\* SEQUENCING IN PROGRESS \*\*\*, in ordered pieces.  
AC019964  
VERSION AC019964.1 GI:6664933  
KEYWORDS HTG: HTGS\_PHASE2.  
SOURCE Drosophila melanogaster.  
ORGANISM Drosophila melanogaster.  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 15707)  
Adams, M. and Venter, J.C.  
Direct Submission  
Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA  
This sequence was identified as CDM:10211545 by the submitter.  
For more information on this record e-mail to fly@celera.com.  
\* NOTE: This is a 'working draft' sequence.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

FEATURES  
source  
1. 15707  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"



Query Match  
Best Local Similarity 92.0%; Score 18.4; DB 1; Length 52486;  
Pred. No. 18;

18770 18869: gap of 100 bp  
18870 19591: contig of 722 bp in length  
19592 19691: gap of 100 bp  
19692 20438: contig of 747 bp in length  
20439 20538: gap of 100 bp  
20539 21224: contig of 686 bp in length  
21225 21324: gap of 100 bp  
21325 22058: contig of 735 bp in length  
22060 22159: gap of 100 bp  
22160 22883: contig of 724 bp in length  
22884 22983: gap of 100 bp  
22984 23702: contig of 719 bp in length  
23703 23802: gap of 100 bp  
23803 24523: contig of 721 bp in length  
24524 24623: gap of 100 bp  
24624 25349: contig of 726 bp in length  
25350 25449: gap of 100 bp  
25450 26140: contig of 691 bp in length  
26141 26240: gap of 100 bp  
26241 26985: contig of 745 bp in length  
26986 27085: gap of 100 bp  
27086 27797: contig of 712 bp in length  
27798 27897: gap of 100 bp  
27898 28618: contig of 721 bp in length  
28619 28718: gap of 100 bp  
28719 29447: contig of 729 bp in length  
29448 29547: gap of 100 bp  
29548 30277: contig of 730 bp in length  
30278 31081: contig of 704 bp in length  
31082 31181: gap of 100 bp  
31182 31918: contig of 737 bp in length  
31919 32018: gap of 100 bp  
32019 32756: contig of 738 bp in length  
32757 32856: gap of 100 bp  
32857 33603: contig of 747 bp in length  
33604 33703: gap of 100 bp  
33704 34414: contig of 711 bp in length  
34415 34514: gap of 100 bp  
34515 35219: contig of 705 bp in length  
35220 35319: gap of 100 bp  
35320 35997: contig of 678 bp in length  
35998 36097: gap of 100 bp  
36098 36801: contig of 704 bp in length  
36802 36901: gap of 100 bp  
36902 37640: contig of 739 bp in length  
37641 37740: gap of 100 bp  
37741 38486: contig of 746 bp in length  
38487 38586: gap of 100 bp  
38587 39331: contig of 745 bp in length  
39332 39431: gap of 100 bp  
39432 40119: contig of 688 bp in length  
40120 40219: gap of 100 bp  
40220 40927: contig of 708 bp in length  
40928 41027: gap of 100 bp  
41028 41741: contig of 714 bp in length  
41742 41841: gap of 100 bp  
41842 42568: contig of 727 bp in length  
42569 42668: gap of 100 bp  
42669 43390: contig of 722 bp in length  
43391 43490: gap of 100 bp  
43491 44216: contig of 726 bp in length  
44217 44316: gap of 100 bp  
44317 45058: contig of 742 bp in length  
45059 45158: gap of 100 bp  
45159 45898: contig of 740 bp in length  
45899 45998: gap of 100 bp  
45999 46710: contig of 712 bp in length  
46711 46810: gap of 100 bp  
46811 47557: contig of 747 bp in length

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TTAGTCTTAGTCTTAGT 20  
Db 28551 TTAGTCTTAGTCTTAGT 28532

RESULT 10  
AC020246/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AC020246  
AC020246.1 GI:6664651  
HTG; HTGS\_PHASE2.  
Drosophila melanogaster  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 56822)  
Adams, M. and Venter, J.C.  
Direct Submission  
Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive,  
Rockville, MD, USA  
This sequence was identified as CDL:10212541 by the submitter.  
For more information on this record e-mail to fly@celera.com.  
\* NOTE: This is a 'working draft' sequence.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
Location/Qualifiers  
1..56822  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"

BASE COUNT 17335 a 11120 c 10936 g 17431 t  
ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 56822;  
Best Local Similarity 95.0%; Pred. No. 18;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TTAGTCTTAGTCTTAGT 20  
Db 8373 TTAGTCTTAGTCTTAGT 8354

RESULT 11  
AC111998/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AC111998  
AC111998.2 GI:21737365  
HTG; HTGS\_PHASE1.  
Norway rat.  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 138256)  
Muzny, D.M., Adams, C., Adio-Oduola, B., All-osman, F.R., Allen, C.,  
Alsbrooks, S.L., Amaralunga, H.C., Are, J.R., Ayele, M., Banks, T.,  
Barbieri, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D.,  
Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,  
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
Carton, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,  
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,  
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,



Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J. H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homi, F., Howard, S., Huber, J., Huliy, S., Hume, J., Jackson, L. E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratoch, J., Kureshi, A., Landry, N., Leal, B., Lewis, L. C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lien, C., Liu, J., Liu, W., Ma, J., Lozado, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M. P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenko, S., Oguh, M., Okunnu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Plinius, E., Pu, L. L., Quiles, M., Ren, Y., Rivas, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoshitari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G. and Gibbs, R.

Unpublished  
Direct Submission  
2 (bases 1 to 138256)  
Worley, K.C.  
Submitted (19-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 138256)  
Worley, K.C.  
Direct Submission  
Submitted (13-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Jul 12, 2002 this sequence version replaced g1:18701948.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: G02S  
Center clone name: CH230-54H8  
----- Summary Statistics  
Sequencing vector: Plasmid  
Chemistry: Dye-terminator Big Dye: 100% of reads  
Assembly program: Phrap: version 0.990329  
Consensus quality: 63539 bases at least Q40  
Consensus quality: 69255 bases at least Q30  
Consensus quality: 73435 bases at least Q20

-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 71 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1037: contig of 1037 bp in length  
\* 1038 1137: gap of unknown length  
\* 1138 2308: contig of 1171 bp in length  
\* 2309 2408: gap of unknown length

2409 3569: contig of 1161 bp in length  
\* 3570 3669: gap of unknown length  
\* 3670 4691: contig of 1022 bp in length  
\* 4692 4791: gap of unknown length  
\* 4792 5888: contig of 1097 bp in length  
\* 5889 5988: gap of unknown length  
\* 5989 7006: contig of 1018 bp in length  
\* 7007 7106: gap of unknown length  
\* 7107 8214: contig of 1108 bp in length  
\* 8215 8314: gap of unknown length  
\* 8315 9514: contig of 1200 bp in length  
\* 9515 9614: gap of unknown length  
\* 9615 10677: contig of 1063 bp in length  
\* 10678 11880: contig of 1103 bp in length  
\* 11881 11980: gap of unknown length  
\* 11981 13176: contig of 1196 bp in length  
\* 13177 13276: gap of unknown length  
\* 13277 14439: contig of 1163 bp in length  
\* 14440 14539: gap of unknown length  
\* 14540 15677: contig of 1138 bp in length  
\* 15678 15777: gap of unknown length  
\* 15778 17258: contig of 1481 bp in length  
\* 17259 17358: gap of unknown length  
\* 17359 19001: contig of 1643 bp in length  
\* 19002 19101: gap of unknown length  
\* 19102 20135: contig of 1034 bp in length  
\* 20136 20235: gap of unknown length  
\* 20236 21365: contig of 1330 bp in length  
\* 21366 21665: gap of unknown length  
\* 21666 23148: contig of 1483 bp in length  
\* 23149 23248: gap of unknown length  
\* 23249 24746: contig of 1498 bp in length  
\* 24747 24846: gap of unknown length  
\* 24847 25905: contig of 1055 bp in length  
\* 25906 26005: gap of unknown length  
\* 26006 27419: contig of 1414 bp in length  
\* 27420 27519: gap of unknown length  
\* 27520 28527: contig of 1008 bp in length  
\* 28528 28627: gap of unknown length  
\* 28628 29768: contig of 1141 bp in length  
\* 29769 29868: gap of unknown length  
\* 29869 31525: contig of 1657 bp in length  
\* 31526 31625: gap of unknown length  
\* 31626 32975: contig of 1350 bp in length  
\* 32976 33075: gap of unknown length  
\* 33076 34670: contig of 1595 bp in length  
\* 34671 34770: gap of unknown length  
\* 34771 36362: contig of 1592 bp in length  
\* 36363 36462: gap of unknown length  
\* 36463 38206: contig of 1744 bp in length  
\* 38207 38306: gap of unknown length  
\* 38307 39464: gap of unknown length  
\* 39465 39564: gap of unknown length  
\* 39565 40936: contig of 1372 bp in length  
\* 40937 41036: gap of unknown length  
\* 41037 43139: contig of 2103 bp in length  
\* 43140 43239: gap of unknown length  
\* 43240 44723: contig of 1484 bp in length  
\* 44724 44823: gap of unknown length  
\* 44824 45944: contig of 1121 bp in length  
\* 45945 46044: gap of unknown length  
\* 46045 47815: contig of 1771 bp in length  
\* 47816 47915: gap of unknown length  
\* 47916 49711: contig of 1796 bp in length  
\* 49712 49811: gap of unknown length  
\* 49812 51328: contig of 1427 bp in length  
\* 51329 51338: gap of unknown length  
\* 51339 52552: contig of 1214 bp in length  
\* 52553 52652: gap of unknown length  
\* 52653 54565: contig of 1913 bp in length  
\* 54566 55938: contig of 1273 bp in length

```

* 55939 56038: gap of unknown length
* 56039 57191: contig of 1153 bp in length
* 57191 57291: gap of unknown length
* 57292 58953: contig of 1662 bp in length
* 58954 59053: gap of unknown length
* 59054 60514: contig of 1461 bp in length
* 60515 60614: gap of unknown length
* 60615 61983: contig of 1369 bp in length
* 61984 62083: gap of unknown length
* 62084 63489: contig of 1406 bp in length
* 63490 63589: gap of unknown length
* 63590 65566: contig of 1977 bp in length
* 63567 65666: gap of unknown length
* 65667 67994: contig of 2328 bp in length
* 67995 68094: gap of unknown length
* 68095 69382: contig of 1288 bp in length
* 69383 69482: gap of unknown length
* 69483 72060: contig of 2578 bp in length
* 72061 72160: gap of unknown length
* 72161 73569: contig of 1409 bp in length
* 73570 73669: gap of unknown length
* 73670 75357: contig of 1688 bp in length
* 75358 75457: gap of unknown length
* 75458 76894: contig of 1437 bp in length
* 76895 76994: gap of unknown length
* 76995 78645: contig of 1651 bp in length
* 78646 78745: gap of unknown length
* 78746 80112: contig of 1367 bp in length
* 80113 80212: gap of unknown length
* 80213 82015: contig of 1803 bp in length

```

```

Query Match      92.08; Score 18.4; DB 1; Length 138256;
Best Local Similarity 95.0%; Pred. No. 20;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 TTAGTCTAGTCTAGT 20
Db 5073 TTACTCTAGTCTAGT 5054

```

```

RESULT 12
AC098044
LOCUS
DEFINITION
AC098044 160054 bp DNA linear HTG 12-JUL-2002
Rattus norvegicus clone CH230-167M23, *** SEQUENCING IN PROGRESS
***, 59 unordered pieces.
AC098044
VERSION AC098044.3 GI:21239825
KEYWORDS HTG: HTGS_PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

```

```

REFERENCE
AUTHORS
1 (bases 1 to 160054)
Wuzy,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunga,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbieri,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouch,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carion,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Doutwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorris,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hoques,M., Holloway,C., Hollins,B.,
Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 160054)
Worley,K.C.
Direct Submission
Submitted (12-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 29, 2002 this sequence version replaced gi:17964025.

```

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GGD8
Center clone name: CH230-167M23

```

```

----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 105329 bases at least Q40
Consensus quality: 112592 bases at least Q30
Consensus quality: 118060 bases at least Q20

```

```

***** NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 59 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1365: contig of 1365 bp in length
* 1366 1465: gap of unknown length
* 1466 2646: contig of 1181 bp in length
* 2647 2746: gap of unknown length
* 2747 3832: contig of 1086 bp in length
* 3833 3932: gap of unknown length
* 3933 5098: contig of 1166 bp in length
* 5099 5198: gap of unknown length
* 5199 6873: contig of 1675 bp in length
* 6874 6973: gap of unknown length
* 6974 8089: contig of 1116 bp in length

```

```

*      8090      8189: gap of unknown length
*      8190      9467: contig of 1278 bp in length
*      9468      9567: gap of unknown length
*      9568      11402: contig of 1835 bp in length
*      11403      11502: gap of unknown length
*      11503      12520: contig of 1018 bp in length
*      12521      12620: gap of unknown length
*      12621      13988: contig of 1278 bp in length
*      13989      15056: gap of unknown length
*      13999      15056: contig of 1058 bp in length
*      15057      15156: gap of unknown length
*      15157      16219: contig of 1063 bp in length
*      16220      16319: gap of unknown length
*      16320      17573: contig of 1254 bp in length
*      17574      17674: gap of unknown length
*      17674      19706: contig of 1933 bp in length
*      19607      21385: gap of unknown length
*      19707      21385: contig of 1679 bp in length
*      21386      21485: gap of unknown length
*      21486      23103: contig of 1618 bp in length
*      23104      23203: gap of unknown length
*      23204      24401: contig of 1198 bp in length
*      24402      24501: gap of unknown length
*      24502      26078: contig of 1577 bp in length
*      26079      26178: gap of unknown length
*      26179      27359: contig of 1181 bp in length
*      27360      27459: gap of unknown length
*      27460      29534: contig of 2075 bp in length
*      29535      29634: gap of unknown length
*      29635      31846: contig of 2212 bp in length
*      31847      31946: gap of unknown length
*      31947      33652: contig of 1706 bp in length
*      33653      33753: gap of unknown length
*      33753      35106: contig of 1354 bp in length
*      35107      35206: gap of unknown length
*      35207      37287: contig of 2081 bp in length
*      37288      37387: gap of unknown length
*      37388      39465: contig of 2078 bp in length
*      39466      39565: gap of unknown length
*      39566      41428: contig of 1863 bp in length
*      41429      41528: gap of unknown length
*      41529      44277: contig of 2749 bp in length
*      44278      44377: gap of unknown length
*      44378      44659: contig of 2092 bp in length
*      44659      46570: gap of unknown length
*      46570      48768: contig of 2199 bp in length
*      48769      48868: gap of unknown length
*      48869      50157: contig of 1289 bp in length
*      50158      50257: gap of unknown length
*      50258      52556: contig of 2299 bp in length
*      52557      52656: gap of unknown length
*      52657      54427: contig of 1771 bp in length
*      54428      54527: gap of unknown length
*      54528      56731: contig of 2204 bp in length
*      56731      56831: gap of unknown length
*      56832      59637: contig of 2806 bp in length
*      59638      59737: gap of unknown length
*      59738      62836: contig of 3099 bp in length
*      62837      62936: gap of unknown length
*      62937      65175: contig of 2239 bp in length
*      65176      65275: gap of unknown length
*      65276      66950: contig of 1675 bp in length
*      66951      67050: gap of unknown length
*      67051      69090: contig of 2040 bp in length
*      69091      69190: gap of unknown length
*      69191      72240: contig of 3050 bp in length
*      72241      72340: gap of unknown length
*      72341      74931: contig of 2591 bp in length
*      74932      75031: gap of unknown length
*      75032      77659: contig of 2628 bp in length
*      77660      77759: gap of unknown length
*      77760      80552: contig of 2793 bp in length
*      80552      80652: gap of unknown length

```

```

*      80653      84775: contig of 4123 bp in length
*      84776      84875: gap of unknown length
*      84876      88523: contig of 3648 bp in length
*      88524      88623: gap of unknown length
*      88624      91755: contig of 3132 bp in length
*      91756      91855: gap of unknown length
*      91856      94622: contig of 2767 bp in length
*      94623      94722: gap of unknown length
*      94723      98734: contig of 4012 bp in length
*      98735      98834: gap of unknown length
*      98835      103639: contig of 4805 bp in length
*      103640      103739: gap of unknown length
*      103740      108154: contig of 4415 bp in length
*      108155      108254: gap of unknown length
*      108255      112250: contig of 3996 bp in length
*      112251      112350: gap of unknown length
*      112351      115708: contig of 3358 bp in length
*      115709      115808: gap of unknown length
*      115809      120163: contig of 4355 bp in length
*      120164      120263: gap of unknown length
*      120264      123938: contig of 3675 bp in length
*      123939      124038: gap of unknown length
*      124039      128955: contig of 4917 bp in length

```

```

Query Match      92.0%  Score 18.4;  DB 1;  Length 160054;
Best Local Similarity 95.0%  Pred. No. 20;
Matches 19;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;

```

```

QY      1  TTAGTCTTAGTCTTAGTCTT 20
      1111111111111111111111
Db      56311 TTAGTCTTAGTCTTAGTCTT 56330

```

```

RESULT 13
AC129283/C
LOCUS
DEFINITION
Rattus norvegicus clone CH230-6117, *** SEQUENCING IN PROGRESS ***
AC129283
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 183218)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsdrooks,S.L., Amaralung,H.C., Are,J.R., Ayala,M., Banks,T.,
Barbata,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durdin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frawitz,P.,
Gabisl,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homsli,F., Howard,S., Huber,J., Huiyk,S., Hume,D., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlssoen,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovac,C.,
Kratovic,D., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtenarge,O., Lien,C., Liu,J., Liu,W., Louised,H.,
Lozado,R.U., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Kapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., Mcleod,M.P., Meador,M., Mel,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,

```

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
Nguyen, N., Nickerson, E., Nwokkenwo, S., Ogih, M., Okunolu, G.,  
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
Peters, L., Pickens, R., Primus, E., Pu, L., L., Quiles, M., Ren, Y.,  
Rives, M., Rojas, A., Rojupokan, I., Rolfe, M., Ruiz, S., Savery, G.,  
Scherer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I.,  
Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H.,  
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
Tansley, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
Umanil, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, O.,  
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S.,  
Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K.,  
Wu, C., Wu, Y., Wu, Y., Zhou, J., Zorrilla, S., Nelson, D.,  
Weinstock, G., and Gibbs, R.  
Direct Submission  
Unpublished  
2 (bases 1 to 183218)  
Morley, K.C.  
Direct Submission  
Submitted (28-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
-----  
Center: Baylor College of Medicine  
Genome Center  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
-----  
Project Information  
Center project name: GBVW  
Center clone name: CH230-6117  
-----  
Summary Statistics  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 127038 bases at least Q40  
Consensus quality: 136608 bases at least Q30  
Consensus quality: 143417 bases at least Q20  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 67 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1  
1002: contig of 1002 bp in length  
1003  
1102: gap of unknown length  
1103  
2859: contig of 1757 bp in length  
2860  
2959: gap of unknown length  
2960  
4118: contig of 1159 bp in length  
4119  
4218: gap of unknown length  
4219  
5370: contig of 1152 bp in length  
5371  
5470: gap of unknown length  
5471  
6980: contig of 1510 bp in length  
6981  
7080: gap of unknown length  
7081  
8087: contig of 1007 bp in length  
8088  
8187: gap of unknown length  
8188  
9322: contig of 1135 bp in length  
9323  
9422: gap of unknown length  
9423  
10530: contig of 1108 bp in length  
10531  
10630: gap of unknown length  
10631  
11677: contig of 1047 bp in length  
11678  
11777: gap of unknown length  
11778  
13579: contig of 1802 bp in length  
13580  
13679: gap of unknown length  
13680  
15247: contig of 1568 bp in length  
15248  
15347: gap of unknown length  
15348  
16962: contig of 1615 bp in length  
16963  
17062: gap of unknown length  
17063  
18375: contig of 1313 bp in length

18376  
18476  
20602  
20702  
22274  
22274  
22374  
24269  
24369  
25604  
25704  
28262  
28362  
28362  
29345  
29645  
30747  
30847  
32515  
32515  
32616  
34231  
34331  
34331  
36381  
36481  
37496  
37596  
40455  
40554  
42501  
42501  
42601  
44711  
44811  
44811  
46467  
46567  
48626  
48725  
48725  
51221  
51221  
51321  
53631  
53731  
53731  
56026  
56125  
57126  
57759  
57859  
59641  
59740  
59741  
61958  
62058  
62058  
64129  
64228  
66519  
66520  
66520  
68462  
68462  
68562  
72071  
72071  
72171  
74415  
74415  
74515  
77259  
77359  
77359  
79617  
79717  
83703  
83803  
83803  
85969  
85969  
86068  
86068  
89794  
89794  
89894  
92509  
92509  
92609  
94913  
94913  
95013  
98757  
98857  
102403  
102404  
18475: gap of unknown length  
20601: contig of 2126 bp in length  
20701: gap of unknown length  
22273: contig of 1572 bp in length  
22373: gap of unknown length  
24268: contig of 1895 bp in length  
24368: gap of unknown length  
25603: contig of 1235 bp in length  
25703: gap of unknown length  
28261: contig of 2558 bp in length  
28361: gap of unknown length  
29544: contig of 1183 bp in length  
29644: gap of unknown length  
30746: contig of 1102 bp in length  
30846: gap of unknown length  
32515: contig of 1669 bp in length  
32615: gap of unknown length  
34230: contig of 1615 bp in length  
34330: gap of unknown length  
36380: contig of 2050 bp in length  
36480: gap of unknown length  
37495: contig of 1015 bp in length  
37595: gap of unknown length  
40454: contig of 2859 bp in length  
40554: gap of unknown length  
42500: contig of 1946 bp in length  
42600: gap of unknown length  
44710: contig of 2110 bp in length  
44810: gap of unknown length  
44810: contig of 1656 bp in length  
46466: gap of unknown length  
46566: gap of unknown length  
48625: contig of 2059 bp in length  
48725: gap of unknown length  
51220: contig of 2495 bp in length  
51320: gap of unknown length  
53630: contig of 2310 bp in length  
53730: gap of unknown length  
56025: contig of 2295 bp in length  
56125: gap of unknown length  
57758: contig of 1633 bp in length  
57858: gap of unknown length  
59640: contig of 1782 bp in length  
59740: gap of unknown length  
61958: contig of 2218 bp in length  
62058: gap of unknown length  
64128: contig of 2070 bp in length  
64228: gap of unknown length  
66519: contig of 2291 bp in length  
66520: gap of unknown length  
68461: contig of 1842 bp in length  
68462: gap of unknown length  
68561: gap of unknown length  
72070: contig of 3509 bp in length  
72071: gap of unknown length  
72170: gap of unknown length  
74414: contig of 2244 bp in length  
74415: gap of unknown length  
77258: contig of 2744 bp in length  
77358: gap of unknown length  
79616: contig of 2258 bp in length  
79716: gap of unknown length  
83702: contig of 3986 bp in length  
83802: gap of unknown length  
85968: contig of 2166 bp in length  
86068: gap of unknown length  
89793: contig of 3725 bp in length  
89893: gap of unknown length  
92508: contig of 2615 bp in length  
92608: gap of unknown length  
94912: contig of 2304 bp in length  
95012: gap of unknown length  
98756: contig of 3744 bp in length  
98856: gap of unknown length  
102403: contig of 3547 bp in length  
102503: gap of unknown length

```

* 102504 105312: contig of 2809 bp in length
* 105313 105412: gap of unknown length
* 105413 109381: contig of 3969 bp in length
* 109382 109481: gap of unknown length
* 109482 113778: contig of 4297 bp in length
* 113779 113878: gap of unknown length
* 113879 116796: contig of 2918 bp in length
* 116797 116896: gap of unknown length
* 116897 119239: contig of 2343 bp in length
* 119240 119339: gap of unknown length
* 119340 124072: contig of 4733 bp in length
* 124073 124172: gap of unknown length
* 124173 128079: contig of 3907 bp in length
* 128080 128179: gap of unknown length
* 128180 131021: contig of 2842 bp in length
* 131022 131121: gap of unknown length

Query Match          92.0%; Score 18.4; DB 1; Length 183218;
Best Local Similarity 95.0%; Pred. No. 20;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 TTAGTCTTACTTCTTACTT 20
DB 33225 TTAGTCTTACTTCTTACTT 33206

```

```

RESULT 14
AC027745 192518 bp DNA linear HTG 04-OCT-2000
LOCUS Homo sapiens clone RP11-535A15, WORKING DRAFT SEQUENCE, 34
DEFINITION unorderd pieces.
ACCESSION AC027745
VERSION AC027745.2 GI:10567968
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 192518)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens, clone RP11-535A15
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 192518)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bede,F.,
Boguski,V., Bouckgeater,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearrellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hages,B., Heatford,A., Horton,L.,
Howard,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Lacroque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGuire,A., McKernan,K., McMeeters,R.,
Meldrum,J., Menus,L., Mihova,T., Miranda,C., Mienna,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Theodore,J., Titrill,A., Travers,M., Tigillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

```

TITLE  
JOURNAL  
COMMENT

Submitted (01-APR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Oct 4, 2000 this sequence version replaced gi:7182564.  
ALL repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research

```

Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L8168
Center clone name: 535_A15
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 171261 bases at least Q40
Consensus quality: 181977 bases at least Q30
Consensus quality: 186186 bases at least Q20
Insert size: 210000; agarose-fp
Insert size: 189218; sum-of-coverage
Quality coverage: 3.1 in Q20 bases; agarose-fp
Quality coverage: 3.4 in Q20 bases; sum-of-coverage
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 34 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1
2775 2874: gap of 100 bp
2775 2874: gap of 100 bp
2875 4195: contig of 1321 bp in length
4196 4295: gap of 100 bp
4296 5880: contig of 1585 bp in length
5881 5980: gap of 100 bp
5981 7755: contig of 1775 bp in length
7756 7855: gap of 100 bp
7856 9023: contig of 1168 bp in length
9024 9123: gap of 100 bp
9124 11085: contig of 1962 bp in length
11086 11185: gap of 100 bp
11186 12742: contig of 1557 bp in length
12743 12842: gap of 100 bp
12843 14362: contig of 1520 bp in length
14363 14462: gap of 100 bp
14463 15891: contig of 1429 bp in length
15892 15991: gap of 100 bp
15992 18140: contig of 2149 bp in length
18141 18240: gap of 100 bp
18241 20252: contig of 2012 bp in length
20253 20352: gap of 100 bp
20353 23974: contig of 3622 bp in length
23975 24074: gap of 100 bp
24075 27028: contig of 2954 bp in length
27029 27128: gap of 100 bp
27129 30517: contig of 3389 bp in length
30518 30617: gap of 100 bp
30618 33387: contig of 2770 bp in length
33388 33487: gap of 100 bp
33488 36476: contig of 2989 bp in length
36477 36576: gap of 100 bp
36577 40395: contig of 3819 bp in length
40396 40495: gap of 100 bp
40496 43638: contig of 3143 bp in length
43639 43738: gap of 100 bp
43739 49735: contig of 5997 bp in length
49736 49835: gap of 100 bp
49836 53882: contig of 4047 bp in length
53883 53982: gap of 100 bp
53983 59666: contig of 5664 bp in length
59667 59766: gap of 100 bp
59767 64782: contig of 5016 bp in length
64783 64882: gap of 100 bp
64883 70725: contig of 5843 bp in length
70726 70825: gap of 100 bp
70826 77243: contig of 6418 bp in length

```

```
* 77244 77343: gap of 100 bp
* 77344 84177: contig of 6834 bp in length
* 84178 84277: gap of 100 bp
* 84278 91428: contig of 7151 bp in length
* 91429 91528: gap of 100 bp
* 91529 101364: contig of 9836 bp in length
* 101365 101464: gap of 100 bp
* 101465 110674: contig of 9210 bp in length
* 110675 110774: gap of 100 bp
* 110775 120859: contig of 10085 bp in length
* 120860 120959: gap of 100 bp
* 120960 131085: contig of 10126 bp in length
* 131086 131185: gap of 100 bp
* 131186 144973: contig of 13788 bp in length
* 144974 145073: gap of 100 bp
* 145074 158411: contig of 13338 bp in length
* 158412 158511: gap of 100 bp
* 158512 173715: contig of 15204 bp in length
* 173716 173815: gap of 100 bp
* 173816 192518: contig of 18703 bp in length.

FEATURES
source
1..192518
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-535A15"
/clone_id="RP11 Human Male BAC"
1..2774
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
2875..4195
/note="assembly_fragment"
4296..5880
/note="assembly_fragment"
5981..7755
/note="assembly_fragment"
7856..9023
/note="assembly_fragment"
9124..11085
/note="assembly_fragment"
11186..12742
/note="assembly_fragment"
12843..14362
/note="assembly_fragment"
14463..15891
/note="assembly_fragment"
15992..18140
/note="assembly_fragment"
18241..20252
/note="assembly_fragment"
20353..23974
/note="assembly_fragment"
24075..27028
/note="assembly_fragment"
27129..30517
/note="assembly_fragment"
30618..33387
/note="assembly_fragment"
33488..36476
/note="assembly_fragment"
36577..40395
/note="assembly_fragment"
40496..43638
/note="assembly_fragment"
43739..49735
/note="assembly_fragment"
49836..53882
/note="assembly_fragment"
53983..59666
/note="assembly_fragment"
59767..64782
/note="assembly_fragment"
64883..70725
/note="assembly_fragment"
```

```
misc_feature /note="assembly_fragment"
70826..77243
misc_feature /note="assembly_fragment"
77344..84177
misc_feature /note="assembly_fragment"
84278..91428
misc_feature /note="assembly_fragment"
91529..101364
misc_feature /note="assembly_fragment"
101365..110674
misc_feature /note="assembly_fragment"
110675..120859
misc_feature /note="assembly_fragment"
120860..131085
misc_feature /note="assembly_fragment"
131086..144973
misc_feature /note="assembly_fragment"
144974..158411
misc_feature /note="assembly_fragment"
158412..173715
misc_feature /note="assembly_fragment"
173716..192518

Query Match 92.0% Score 18.4; DB 1; Length 192518;
Best Local Similarity 95.0%; Pred. No. 21;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTAGTCTTCTTACTTACTT 20
Db 164920 TTAGTCTTCTTACTTACTT 164939

RESULT 15
AC099584/c
LOCUS
DEFINITION Mus musculus clone RP23-1A12, WORKING DRAFT SEQUENCE, 12 unordered
pieces.
ACCESSION AC099584.1 GI:16946033
VERSION AC099584.1
KEYWORDS HTG; HTGS-PHASE1; HTGS-DRAFT; HTGS-FULLTOP.
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 217395)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Bouckgeater,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Illiev,I., Johnson,R.,
Jones,C., Kamat,A., Karats,A., Kells,C., LaRoque,K.,
Lamazares,R., Landers,T., Lehoczek,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrum,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green,P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L17136
Center clone name: L_A12
----- Summary Statistics
```

Sequencing vector: Plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 214412 bases at least Q40  
 Consensus quality: 215767 bases at least Q30  
 Consensus quality: 216139 bases at least Q20  
 Insert size: 212000; agarose-fp  
 Insert size: 216295; sum-of-contigs  
 Quality coverage: 9.6 in Q20 bases; agarose-fp  
 Quality coverage: 9.4 in Q20 bases; sum-of-contigs

-----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 12 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

```

1      51897: contig of 51897 bp in length
*      51898 51997: gap of 100 bp
*      51998 54376: contig of 2379 bp in length
*      54377 54476: gap of 100 bp
*      54477 58077: contig of 3601 bp in length
*      58078 58177: gap of 100 bp
*      58178 61196: contig of 3019 bp in length
*      61197 61296: gap of 100 bp
*      61297 66269: contig of 4973 bp in length
*      66270 66369: gap of 100 bp
*      66370 74862: contig of 8493 bp in length
*      74863 74962: gap of 100 bp
*      74963 86419: contig of 11457 bp in length
*      86420 86519: gap of 100 bp
*      86520 108007: contig of 21488 bp in length
*      108008 108107: gap of 100 bp
*      108108 132537: contig of 24430 bp in length
*      132538 132637: gap of 100 bp
*      132638 159982: contig of 27345 bp in length
*      159983 160082: gap of 100 bp
*      160083 206764: contig of 46682 bp in length
*      206765 206864: gap of 100 bp
*      206865 217395: contig of 10531 bp in length.

```

## FEATURES

## source

```

1. .217395
  /organism="Mus musculus"
  /db_xref="taxon:10090"
  /clone_lib="RPCI-23 Female Mouse BAC"
  1. .51897
    /note="assembly-fragment"
    clone_end:SP6
    vector_side:left"
misc_feature
  51998. .54376
    /note="assembly-fragment"
misc_feature
  54477. .58077
    /note="assembly-fragment"
misc_feature
  58178. .61196
    /note="assembly-fragment"
misc_feature
  61297. .66269
    /note="assembly-fragment"
misc_feature
  66370. .74862
    /note="assembly-fragment"
misc_feature
  74963. .86419
    /note="assembly-fragment"
misc_feature
  86520. .108007
    /note="assembly-fragment"
misc_feature
  108108. .132537
    /note="assembly-fragment"
misc_feature
  132638. .159982
    /note="assembly-fragment"
misc_feature
  160083. .206764
    /note="assembly-fragment"
misc_feature
  206865. .217395
    /note="assembly-fragment"

```

```

/note="assembly-fragment"
clone_end:r7
vector_side:right"
BASE COUNT      63883 a 44235 c 45035 g 63142 t 1100 others
ORIGIN

```

```

Query Match      92.0%; Score 18.4; DB 1; Length 217395;
Best Local Similarity 95.0%; Pred No. 21;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 TTAGTCTTAGTCTTAGTT 20

```

```

Db 135236 TTAGTCTTAGTCTTAGTT 135217

```

```

Search completed: April 6, 2003, 11:50:25
Job time : 463.5 secs

```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 10:34:07 ; Search time 243 Seconds

(without alignments)  
1069.066 Million cell updates/sec

Title: US-09-980-265-11

Perfect score: 20

Sequence: 1 ttagcttctagcttctagct 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1994485 seqs, 6494577260 residues

Total number of hits satisfying chosen parameters: 3988970

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl\_NOHTG\*

1: gb\_ba:\*

2: gb\_in:\*

3: gb\_cm:\*

4: gb\_ov:\*

5: gb\_pat:\*

6: gb\_ph:\*

7: gb\_pl:\*

8: gb\_pr:\*

9: gb\_ro:\*

10: gb\_sts:\*

11: gb\_sy:\*

12: gb\_un:\*

13: gb\_vl:\*

14: em\_ba:\*

15: em\_fun:\*

16: em\_hum:\*

17: em\_in:\*

18: em\_mu:\*

19: em\_om:\*

20: em\_or:\*

21: em\_ov:\*

22: em\_pat:\*

23: em\_ph:\*

24: em\_pl:\*

25: em\_ro:\*

26: em\_sts:\*

27: em\_un:\*

28: em\_vl:\*

29: em\_htg\_hum:\*

30: em\_htg\_inv:\*

31: em\_htg\_other:\*

32: em\_htg\_mus:\*

33: em\_htg\_pln:\*

34: em\_htg\_rtd:\*

35: em\_htg\_mam:\*

36: em\_htg\_vrt:\*

37: em\_sy:\*

38: em\_htgo\_hum:\*

39: em\_htgo\_mus:\*

40: em\_htgo\_other:\*

and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	20	100.0	20	5	AX057377	AX057377 Sequence
2	20	100.0	65	5	AX485115	AX485115 Sequence
3	19	95.0	56820	7	DM87C10	ALI21804 Drosophila
4	19	95.0	91268	2	AP004895	AP004895 Lotus_jap
5	19	95.0	191342	2	AC104140	AC104140 Drosophila
6	19	95.0	225065	9	AC121835	AC121835 Mus muscu
7	19	95.0	289090	2	AE003424	AE003424 Drosophila
8	18.4	92.0	20	5	AX057369	AX057369 Sequence
9	18.4	92.0	20	5	AX057373	AX057373 Sequence
10	18.4	92.0	20	5	AX057375	AX057375 Sequence
11	18.4	92.0	20	5	AX057376	AX057376 Sequence
12	18.4	92.0	20	5	AX057379	AX057379 Sequence
13	18.4	92.0	20	5	AX057381	AX057381 Sequence
14	18.4	92.0	36942	8	AC093782	AC093782 Homo sapi
15	18.4	92.0	61372	8	AC092116	AC092116 Homo sapi
16	18.4	92.0	71321	7	AP004519	AP004519 Lotus_jap
17	18.4	92.0	78249	7	AP004529	AP004529 Lotus_jap
18	18.4	92.0	78574	7	AP004536	AP004536 Lotus_jap
19	18.4	92.0	87900	7	AP004475	AP004475 Lotus_jap
20	18.4	92.0	100810	7	AF411804	AF411804 Lycopersi
21	18.4	92.0	100810	7	AP004962	AP004962 Lotus_jap
22	18.4	92.0	101270	7	AP004942	AP004942 Lotus_jap
23	18.4	92.0	115046	7	AP004907	AP004907 Lotus_jap
24	18.4	92.0	151130	9	AL139093	AL139093 Human DNA
25	18.4	92.0	151395	8	AC127432	AC127432 Mus muscu
26	18.4	92.0	152492	8	AC079863	AC079863 Homo sapi
27	18.4	92.0	152545	2	AC010922	AC010922 Drosophila
28	18.4	92.0	158801	8	AC079118	AC079118 Homo sapi
29	18.4	92.0	163464	8	AC007462	AC007462 Homo sapi
30	18.4	92.0	164826	2	AC099038	AC099038 Drosophila
31	18.4	92.0	170356	2	AC010558	AC010558 Drosophila
32	18.4	92.0	175867	2	AC007976	AC007976 Drosophila
33	18.4	92.0	179016	2	AC018489	AC018489 Drosophila
34	18.4	92.0	190014	7	AP000836	AP000836 Drosophila
35	18.4	92.0	203831	8	AC093295	AC093295 Homo sapi
36	18.4	92.0	244164	2	AE003548	AE003548 Drosophila
37	18.4	92.0	300474	2	AE003503	AE003503 Drosophila
38	18.4	92.0	301783	2	AE003499	AE003499 Drosophila
39	18.4	92.0	324484	2	AE003584	AE003584 Drosophila
40	18	90.0	94052	7	AP004903	AP004903 Lotus_jap
41	18	90.0	97268	7	AP004546	AP004546 Lotus_jap
42	18	90.0	97268	7	AP004546	AP004546 Lotus_jap
43	18	90.0	106023	7	AP004467	AP004467 Lotus_jap
44	18	90.0	123336	8	AC003075	AC003075 Homo sapi
45	18	90.0	131741	7	AP004916	AP004916 Lotus_jap

# ALIGNMENTS

RESULT 1  
AX057377  
LOCUS AX057377 20 bp DNA  
DEFINITION Sequence 11 from Patent WO0075304.  
ACCESSION AX057377  
VERSION AX057377.1 GI:12310118  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct.  
synthetic construct.  
artificial sequences.  
REFERENCE  
AUTHORS Bachy,M., Sodoyer,R. and Tranmy,E.  
TITLE Immunostimulant oligonucleotide  
JOURNAL Patent: WO 0075304-A II 14-DEC-2000;  
Aventis Pasteur (FR)  
FEATURES  
Location/Qualifiers

source 1. .20  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="oligonucleotide"

BASE COUNT 3 a 2 c 3 g 12 t  
ORIGIN

Query Match 100.0%; Score 20; DB 5; Length 20;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTACTTCTTAGTCTTAGTT 20  
|||||  
Db 1 TTACTTCTTAGTCTTAGTT 20

RESULT 2  
AX485115/c AX485115 65 bp DNA linear PAT 16-AUG-2002  
DEFINITION Sequence 2415 from Patent WO02053728.  
ACCESSION AX485115  
VERSION AX485115.1 GI:22319399  
KEYWORDS  
SOURCE  
ORGANISM  
Candida albicans.  
Candida albicans.  
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; mitosporic Saccharomycetales; Candida.

REFERENCE 1  
AUTHORS Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.  
TITLE Gene disruption methodologies for drug target discovery  
JOURNAL Patent: WO 02053728-A 2415 11-JUL-2002;  
Elitra Pharmaceuticals, Inc. (US)  
FEATURES  
source 1. .65  
/organism="Candida albicans"  
/db\_xref="taxon:5476"

BASE COUNT 37 a 6 c 16 g 6 t  
ORIGIN

Query Match 100.0%; Score 20; DB 5; Length 65;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTACTTCTTAGTCTTAGTT 20  
|||||  
Db 46 TTACTTCTTAGTCTTAGTT 27

RESULT 3  
DMB7C10 56820 bp DNA linear INV 07-FEB-2000  
LOCUS Drosophila melanogaster BAC clone BACR7C10.  
DEFINITION AL121804  
ACCESSION AL121804  
VERSION AL121804.2 GI:6691820  
KEYWORDS  
SOURCE  
ORGANISM  
Drosophila melanogaster.  
Drosophila melanogaster.  
Eukaryota; Metazoa; Insecta; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE 1 (bases 1 to 56820)  
AUTHORS Ferraz,C., Vidal,S., Brun,C., Bucheton,A. and Demaille,J.C.  
TITLE Sequencing the distal X chromosome of Drosophila melanogaster  
JOURNAL Institut de Genetique Humaine -UPR 1142- CNRS, 141,rue de la  
REMARK Cardonille, 34396 Montpellier Cedex 5, France  
2 (bases 1 to 56820)  
REFERENCE Benos,P.  
AUTHORS Direct Submission  
TITLE Submitted (06-FEB-2000) European Drosophila Genome Sequencing  
JOURNAL Consortium  
COMMENT On Jan 12, 2000 this sequence version replaced g1:6018822.  
Sequence submitted by Takis Benos, EMBL Outstation - The EBI,

hinxton, Cambridge, CB10 1SD, U.K.  
E-mail: benos@ebi.ac.uk on behalf of the European Drosophila Genome  
Sequencing Consortium. For further information see the European  
Drosophila Genome Sequencing Consortium's web site:  
<http://edgp.ebi.ac.uk/>.  
The syntax for the representation of annotation used in this record  
is documented at:  
[ftp://ftp.ebi.ac.uk/pub/databases/edgp/sequence\\_annotation\\_README](ftp://ftp.ebi.ac.uk/pub/databases/edgp/sequence_annotation_README)  
Coding sequences are predicted from computer analysis, using both  
gene and CDS prediction programs and matches to other sequences.  
These predictions and matches have been evaluated by the annotators  
and may have been refined by hand (in which case a GeneFinder  
prediction will have no score. The annotators have also used their  
judgement on what matches to represent in this record. A far more  
complete annotation record is available from FlyBase  
(<http://flybase.bio.indiana.edu/>) through the FlyBase Annotation  
Object linked by the db\_xref qualifier in the feature table.  
IMPORTANT: This sequence is NOT necessarily the entire insert of  
clone BACR7C10. It may be shorter, since we are minimising the  
overlap between clones to 100 bases, by trimming them. Sequence in  
absolute orientation with respect to chromosome clone BACR7C10  
overlaps to the left with clone BACR25B3 clone BACR7C10 overlaps to  
the right with clone 155E2 The true right end of clone BACR7C10 is  
unknown  
clone-BACR7C10; Contig ID=1; Length=56820; Status=finished.

FEATURES  
source  
1. .56820  
Location/Qualifiers  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/clone="BAC BACR7C10"  
join(41. .610,672. .1487)  
/gene="EG:BACR7C10.3"  
/note="Fgdn0004643:mt(1)15"  
join(<41. .610,672. .1487)  
/gene="EG:BACR7C10.3"  
/codon\_start=1  
/protein\_id="CAB5854.1"  
/db\_xref="GI:6691821"  
/translation="MRPVADVSEADGGTVRLSLSTATKEPSSAHVRPNKQVLN  
LRLHLTLGAINCSVSDQHVFGIIGDHVKDKMLKLIVBCLIPAVESESEVQSTL  
CEDVALEOLIVDSFTINPODRLALQDFEYKTYLRNRYRLTARLTIORDLD  
MVLAVANNNSAEVANDPFLPRCMISKSDQFVKMDRILOPTDKLGDDEADPIGV  
ISTMLRTYINEVPRVHKRLLESTPDQAVFPHNNCMFTFHWAGHAKNGJESLAIAKT  
LQATGQGHFRVQVYDOSTLMGTMOEPEFESTITLDSGPLKLYRQCLROLELNKWA  
NVLPEYVNAATFCGLINTFVAILIRFVLRHSIAQMACLSLDIVLQRAPLTRE  
PNEVVOVLSWLKLOQLKAMLNASIMEITELMGDVGPLASYSDEIKHIALFQDT  
DMRAKAITQIV"  
complement(join(1784. .2154,2220. .2682,2928. .3061,  
3129. .3370,3429. .3625))  
/gene="EG:BACR7C10.4"  
complement(join(1784. .2154,2220. .2682,2928. .3061,  
3129. .3370,3429. .3625))  
/gene="EG:BACR7C10.4"  
/note="EG:BACR7C10.4"  
/version="084", score="75.04",  
/prediction-method="genscan", version="1.0",  
score="178.97",  
/match-desc="2INC-FINGER PROTEIN BOP", species="Mus  
musculus (Mouse)", ranges=(query:3262. .3366,  
target:SPPREMBL::p97443:95. .61, score:"109.00"),  
(query:2374. .2508, target:SPPREMBL::p97443:291. .247,  
score:"98.00"), (query:1844. .2146,  
target:SPPREMBL::p97443:431. .331, score:"137.00"),  
method:"blastx", version:"1.4.9")"  
/codon\_start=1  
/protein\_id="CAB58065.1"  
/db\_xref="GI:6018870"  
/translation="MATRYGSSRSATSTNTASCSKSNKLNKPNAPQIKRGRIILTE  
KPPAFVLDKSOYRLRCDNCEATKVLKSCRCRVSYCHRSQOMAGOHHECPFLIK  
VHPRVVDAKMRICRLILREHGDLIRGYEYRKRFDMLSHVAELKNDPMRLKH  
LDSLHVLVDMMAESPSPVNPKNALDELITNGFNILADAMNSIAVALYGVST  
DHSQCNAAVATFEGNELHVAHAIEDMCLMSKFFIYIDLINPEORLDLKHYYFL  
CVSKCTDAKESKEMIALCPNRCAGISVDNNNCPRODAGISPLRNANFNAMLT

```

gene
    repeat_unit
        gene
            complement(5405..12456)
            /gene="EG:BACR7C10.5"
            /note="FBgn00101302;Burdock"
            complement(5405..12456)
            /gene="FBgn00101302;Burdock"
            /function="transposon"
            complement(join(13393..13718,13826..13996,14073..14169,14175..14489))
            /gene="EG:BACR7C10.6"
            complement(join(13393..13718,13826..13996,14073..14169,14175..14489))
            /gene="EG:BACR7C10.6"
            /note="EG:BACR7C10.6"
            /prediction=(method:"genscan", version:"1.0")
            /match=(desc:"TETRASPAN NET-4", species:"Homo sapiens (Human)", ranges:(query:14274..14375, target:SPTREMBL::060746:44..11, score:"96.00"), (query:14071..14154, target:SPTREMBL::060746:109..82, score:"90.00"), (query:13854..13943, target:SPTREMBL::060746:155..126, score:"97.00"), (query:13402..13713, target:SPTREMBL::060746:268..165, score:"159.00"), method:"blastx", version:"1.4.9"), /match=(desc:"CELL SURFACE GLYCOPROTEIN A15 (T-CELL ACUTE LYMPHOBLASTIC LEUKEMIA ASSOCIATED ANTIGEN 1) (TALL-1) (MEMBRANE COMPONENT, X CHROMOSOME, SURFACE MARKER 1)", species:"Homo sapiens (Human)", ranges:(query:14289..14375, target:SWISS-PROT::P41732:34..6, score:"75.00"), (query:14065..14145, target:SWISS-PROT::P41732:103..77, score:"53.00"), (query:13845..13928, target:SWISS-PROT::P41732:150..123, score:"51.00"), (query:13636..13695, target:SWISS-PROT::P41732:175..156, score:"56.00"), (query:13444..13548, target:SWISS-PROT::P41732:229..195, score:"73.00"), method:"blastx", version:"1.4.9")
            /match=(desc:"GH15125.5prime GH Drosophila melanogaster head port2 Drosophila melanogaster cDNA clone GH15125 5prime, mRNA sequence", species:"Drosophila melanogaster (fruit fly)", ranges:(query:14153..14694, target:EMBL::A1238968:542..1, score:"2710.00"), (query:14153..14694, target:EMBL::A1238968:542..1, score:"2710.00"), (query:14081..14154, target:EMBL::A1238968:615..542, score:"370.00"), method:"blastn", version:"1.4.9")
            /codon_start=1
            /protein_id="CAB65855.1"
            /db_xref="GI:6691822"
            /translation="MSNRYOGAGIGVGMAGRGSGIEVHEWPHNHTYVSCVK YMTILNPFVFLFGGLIGVYAFRDKMEKANGSVLENTYDFLNSLYMTLAGTY IFTLLRLGALRENTFLKTYSMCLLEFLLEMAIATVCFPOPMNTEFKTKTKI IHSYRDPDIDNFIIDFAQEFKCGLSNGSGYDMSKSNFSCSVKCPVSCCI NATDSSGLVIMCGYGVONAPVEATKLTMTSCIEIVRWAEHNILVIAGNALGIA LIOLVLYLAKTIEGQIELOKSRMLA"
            join(15266..15549,15781..16268,16335..16675)
            /gene="EG:BACR7C10.1"
            join(15266..15549,15781..16268,16335..16675)
            /gene="EG:BACR7C10.1"
            /note="EG:BACR7C10.1"
            /prediction=(method:"genscan", version:"1.0", score:"171.47")
            /prediction=(method:"genefinder", version:"084")
            /match=(desc:"LD08769.5prime LD Drosophila melanogaster embryo Bluescript Drosophila melanogaster cDNA clone LD08769 5prime, mRNA sequence", species:"Drosophila melanogaster (fruit fly)", ranges:(query:14795..15334, target:EMBL::AA390535:1..540, score:"2691.00"), method:"blastn", version:"1.4.9"), /match=(desc:"SD04409.5prime SD Drosophila melanogaster Schneider L2 cell culture port2 Drosophila melanogaster cDNA clone SD04409 5prime, mRNA sequence", species:"Drosophila melanogaster (fruit fly)",

```

```

ranges:(query:15095..15549, target:EMBL::A1532791:1..455, score:"2230.00"), (query:15778..15839, target:EMBL::A1532791:453..514, score:"292.00"), method:"blastn", version:"1.4.9"), /match=(desc:"HYPOTHETICAL 46.2 KD PROTEIN", species:"Mus musculus (Mouse)", ranges:(query:15410..15550, target:SPTREMBL::0922E9:52..98, score:"1139.00"), (query:15842..16282, target:SPTREMBL::0922E9:123..269, score:"289.00"), method:"blastx", version:"1.4.9"), /codon_start=1
/protein_id="CAB65856.1"
/db_xref="GI:6691823"
/translation="MNTILRLIVFALDPLGGRFLIRPAVNLGMVYDRVRSKADK VGVRELVRLGLIAFVAVLITLAVEMVYAAFYVMPAISHTRPVHOFKTCLETST PCTPRAHVSLTKKQOLMVGQAKVIVIMPSPOLIEIGMFVACEMDQSMRL GHSCRSAMRRSPRLIRNISTWVSLPLVIGMKKEFOQVPEYFISRYLEEQHRTIV YVEIOSKIQRTYVTLHIVADFGRTIMEVMPVLSAISTNLFLLVFLLSWTH WSDAKWLHSVQIKYARLRKSLPEGVHSKASLRLDDDLVAYSDKSDIADVGDTLIS DVDADDLVLVKKRSRSGKRESDALRKPTKRTTADH"
complement(join(17148..18350,18385..18468,18549..19451))
/gene="EG:BACR7C10.7"
complement(join(17148..18350,18385..18468,18549..19451))
/gene="EG:BACR7C10.7"
/note="EG:BACR7C10.7"
/prediction=(method:"genscan", version:"1.0", score:"356.84")
/prediction=(method:"genefinder", version:"084")
/match=(desc:"POLYCYSTIC KIDNEY DISEASE AND RECEPTOR FOR EGG JELLY RELATED PROTEIN", species:"Homo sapiens (Human)", ranges:(query:18733..18436, target:SPTREMBL::095850:1928..1901, score:"71.00"),
Query Match
Best Local Similarity 95.0% Score 19; DB 2; Length 56820;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTAGTCTTACTTCTTACT 19
DB 34495 TTAGTCTTACTTCTTACT 34513
RESULT 4
AP004895/C 91268 bp DNA linear PLN 19-JUL-2002
LOCUS Lotusc japonicus genomic DNA, chromosome 4, clone:LJT02D01, TM0039,
DEFINITION complete sequence.
ACCESSION AP004895
VERSION AP004895.1 GI:21907911
KEYWORDS HTG.
SOURCE Lotusc japonicus
ORGANISM Lotusc japonicus DNA, clone_ltb:LJT library clone:LJT02D01.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotusc.
REFERENCE
1 Kaneko,T., Nakamura,Y., Asamizu,E., Kato,T., Sato,S. and Tabata,S. Structural Analysis of a Lotusc japonicus Genome. I. Sequence Features and Mapping of Sixty-six TAC clones which cover the 6.7 Mb Regions of the Genome
2 (bases 1 to 91268)
JOURNAL Unpublished
REFERENCE Nakamura,Y.
AUTHORS Direct Submission
TITLE Submitted (26-Mar-2002) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research, 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:yn@kazusa.or.jp, URL:http://www.kazusa.or.jp, Tel:81-438-52-3935, Fax:81-438-52-3934)
FEATURES
SOURCE Location/Qualifiers
1. 91268 /organism="Lotusc japonicus"
/db_xref="taxon:34305"

```

BASE COUNT 27763 a 20258 c 17608 g 25639 t  
ORIGIN  
Query Match 95.0%; Score 19; DB 7; Length 91268;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 2 TACTCTTACTCTTACTT 20  
Db 61908 TAGTCTTACTCTTACTT 61890  
RESULT 5  
AC104140  
LOCUS  
DEFINITION AC104140 191342 bp DNA linear INV-24-JUL-2002  
Drosophila melanogaster X BAC RP98-9H13 (Roswell Park Cancer  
AC104140  
INSTITUTE Drosophila BAC Library) complete sequence.  
AC104140.5 GI:17940805  
HTG.  
Drosophila melanogaster.  
Drosophila melanogaster.  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 191342)  
Wuzny,D., Scherer,S., Adams,M.D., Holt,R.A., Evans,C.A.,  
Gocayne,J.D., Taber,P., Williamson,A.H., Homsl,F.H.,  
Dugan-Rocha,S.D., Sodergren,E.S., Hodgson,A.H., Chen,R.C.,  
Ayele,M.A., Scott,G.S., Worley,K.W., Amandides,P.G., Brandon,R.C.,  
Rogers,Y., An,H., Baldwin,D., Beeson,K.Y., Brown,M., Buhay,C.,  
Busan,D.A., Center,A., Chen,G., Chen,Z., Clerc-Blankenburg,K.,  
Davenport,L.B., Dietz,S.M., Ding,Y., Dodson,K., Domp,L.E.,  
Draper,H., Emery-Cohen,A., Ferrier,S., Garg,N.D.S., Houck,J.,  
Hostin,D., Howland,T.J., Hume,J., Ibegwam,C., Jallil,M., Kovar,C.,  
Liu,M., Mattei,B., McIntosh,T.C., Morgan,M., Moy,M., Murphy,B.,  
Nelson,K.A., Ndassa,Y., Nguyen,N., Perez,L., Pittman,G.S., Puri,V.,  
Schaefer,F., Shen,H., Strong,R., Tector,C., Wang,O., Williams,S.M.,  
Xiang,J., Zaveri,J.S., Zhou,J., Zorrilla,S., Smith,H.O.,  
Wheeler,D., Weinstein,G., Gibbs,R. and Venter,J.C.  
Direct Submission  
Unpublished  
2 (bases 1 to 191342)  
Worley,K.C.  
Direct Submission  
Submitted (05-DEC-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 191342)  
Worley,K.C., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
Alstbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbara,J.,  
Benton,J., Bimige,K., Blankenburg,K., Bonnin,D., Bouck,J.,  
Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,  
Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F.,  
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,  
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,  
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,  
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,  
Deen,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,  
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,  
Elhaj,C., Escotto,M., Falls,T., Ferrante,D., Flagg,N., Ford,J.,  
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,  
Garrza,N., Gill,R., Gorrell,H., Guevara,W., Gunaratne,P., Hale,S.,  
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,  
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,  
Hollins,B., Homsl,F., Howard,S., Huber,J., Huliyk,S., Hume,J.,  
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,  
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,  
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,

REFERENCE  
AUTHORS  
JOURNAL

Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,M.,  
Loui-seged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,  
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,  
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,  
Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabab,K.,  
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,  
Nguyen,A., Nguyen,N., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,  
Ogih,M., Okunou,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,  
Peery,J., Perez,L., Peters,L., Pickens,R., Plums,E., Pu,L.L.,  
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubenkan,I., Rolfe,M.,  
Rulz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,  
Stinson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,  
Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K.,  
Tang,H., Tansey,J., Taylor,C., Taylor,T., Tellford,B., Thomas,N.,  
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,  
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,  
Washington,S., Williams,G., Williamson,A., Wleczek,R., Wooden,S.,  
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstein,G., and Gibbs,R.  
Direct Submission  
Submitted (20-DEC-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
4 (bases 1 to 191342)  
Worley,K.C., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
Alstbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbara,J.,  
Benton,J., Bimige,K., Blankenburg,K., Bonnin,D., Bouck,J.,  
Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,  
Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F.,  
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,  
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,  
Coyle,M.D., Dathorne,S.R., David,R., David,R., Davila,M.L., Davis,C.,  
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,  
Deen,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,  
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,  
Elhaj,C., Escotto,M., Falls,T., Ferrante,D., Flagg,N., Ford,J.,  
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,  
Garrza,N., Gill,R., Gorrell,H., Guevara,W., Gunaratne,P., Hale,S.,  
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,  
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,  
Hollins,B., Homsl,F., Howard,S., Huber,J., Huliyk,S., Hume,J.,  
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,  
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,  
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,  
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,M.,  
Loui-seged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,  
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,  
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,  
Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabab,K.,  
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,  
Nguyen,A., Nguyen,N., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,  
Ogih,M., Okunou,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,  
Peery,J., Perez,L., Peters,L., Pickens,R., Plums,E., Pu,L.L.,  
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubenkan,I., Rolfe,M.,  
Rulz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,  
Stinson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,  
Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K.,  
Tang,H., Tansey,J., Taylor,C., Taylor,T., Tellford,B., Thomas,N.,  
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,  
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,  
Washington,S., Williams,G., Williamson,A., Wleczek,R., Wooden,S.,  
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstein,G., and Gibbs,R.  
Direct Submission  
Submitted (27-DEC-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
5 (bases 1 to 191342)  
Worley,K.C., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
Alstbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbara,J.,  
Benton,J., Bimige,K., Blankenburg,K., Bonnin,D., Bouck,J.,  
Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,  
Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F.,

REFERENCE  
AUTHORS  
JOURNAL

Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,M.,  
Loui-seged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,  
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,  
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,  
Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabab,K.,  
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,  
Nguyen,A., Nguyen,N., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,  
Ogih,M., Okunou,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,  
Peery,J., Perez,L., Peters,L., Pickens,R., Plums,E., Pu,L.L.,  
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubenkan,I., Rolfe,M.,  
Rulz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,  
Stinson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,  
Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K.,  
Tang,H., Tansey,J., Taylor,C., Taylor,T., Tellford,B., Thomas,N.,  
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,  
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,  
Washington,S., Williams,G., Williamson,A., Wleczek,R., Wooden,S.,  
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstein,G., and Gibbs,R.  
Direct Submission  
Submitted (27-DEC-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
5 (bases 1 to 191342)  
Worley,K.C., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
Alstbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbara,J.,  
Benton,J., Bimige,K., Blankenburg,K., Bonnin,D., Bouck,J.,  
Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,  
Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F.,

Carter, M., Cavazos, S. R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C. D., Cox, C., Coyle, M. D., Dathorne, S. R., David, R., Davila, M. L., Davis, C., Davy-Carroll, L., Dederich, D. A., Delaney, K. R., Delgado, O., Dem, A. L., Ding, Y., Dinh, H. H., Doulatbait, K. J., Draper, H., Dugan-Rocha, S., Durbin, K. J., Ebnhart, C., Edgar, D., Edwards, C. C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J. H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, U., Huik, S., Hume, J., Jackson, L. E., Jacobson, B., Jia, Y., Johnson, R., Jollivel, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L. C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H., Lozano, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maneshwari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhinney, E., McLeod, M. P., Meador, M., Mel, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L. L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojuben, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shooshari, N., Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Tellford, B., Thomas, R., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalob, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

Direct Submission  
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Dec 20, 2001 this sequence version replaced gi:17736993.  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email [gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

TITLE  
JOURNAL  
COMMENT

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

## ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base

Query Match 95.0%; Score 19; DB 2; Length 191342;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTAGT 19  
|||||  
Db 143942 TTAGTCTTAGTCTTAGT 143960

RESULT 6  
AC121835  
LOCUS  
DEFINITION  
AC121835  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AC121835 225065 bp DNA linear ROD 13-AUG-2002  
Mus musculus chromosome 8 clone RP24-72L18, complete sequence.  
AC121835  
AC121835.2 GI:22213523  
HTG.  
house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 225065)  
McPherson, J.D. and Waterston, R.H.  
The sequence of Mus musculus clone  
Unpublished  
2 (bases 1 to 225065)  
McPherson, J.D. and Waterston, R.H.  
Direct Submission  
Submitted (21-MAY-2002) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
3 (bases 1 to 225065)  
McPherson, J.D. and Waterston, R.H.  
Direct Submission  
Submitted (13-AUG-2002) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
On Aug 13, 2002 this sequence version replaced gi:21039928.

COMMENT

FEATURES  
source  
location/Qualifiers  
1..225065  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosome="8"  
/clone="RP24-72L18"

BASE COUNT 60052 a 50559 c 50972 g 63482 t

ORIGIN

Query Match 95.0%; Score 19; DB 9; Length 225065;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTAGT 19  
|||||  
Db 148339 TTAGTCTTAGTCTTAGT 148357

RESULT 7  
AE003424  
LOCUS  
DEFINITION  
AC121835  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AE003424 289090 bp DNA linear INV 04-OCT-2000  
Drosophila melanogaster genomic scaffold 142000013386054 section 8  
of 35, complete sequence.  
AE003424 AE002566  
AE003424.1 GI:7290316  
HTG.  
Drosophila melanogaster.  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera: Endopterygota: Diptera: Brachycera: Muscomorpha: Ephydroidea: Drosophilidae: Drosophila.

REFERENCE  
AUTHORS  
1 (bases 1 to 289090)  
Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F., George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N., Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X., Brandon,R.C., Rogers,Y.H., Blazej,R.G., Champe,M., Pfeiffer,B.D., Man,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor,M.K., Doyle,J.F., Agbayani,A., An,H.J., Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Basu,A., Baxendale,J., Bayraktaroglu,T., Beasley,E.M., Beeson,K.Y., Benos,P.V., Bernier,B.P., Bhandari,D., Bolshakov,S., Botkova,D., Botchan,M.R., Bouck,J., Brokstein,P., Brotlier,P., Burks,K.C., Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I., Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M., Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C., Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferrieres,S., Fleischmann,W., Foster,C., Gabrielian,A.E., Garg,N.S., Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z., Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J., Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J., Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke-Z., Krasits,S., Kulp,D., Lai,Z., Lasko,P., Del,Y., Levitsky,A.A., Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C., McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C., Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L., Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K., Nusser,D.R., Pacle,J.M., Palazzolo,M., Pittman,G.S., Pan,S., Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K., Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I., Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C., Stapleton,M., Strong,R., Sun,E., Svirskas,R., Tector,C., Turner,R., Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A., Weinstock,G.M., Weissbach,J., Williams,S.M., Woodage,T., Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yen,R.F., Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H., Zhong,F.N., Zhong,W., Zhou,S., Zhu,X., Zhu,X., Smith,H.O., Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.

THE genome sequence of *Drosophila melanogaster*  
Science 287 (5461), 2185-2195 (2000)

TITLE  
JOURNAL  
MEDLINE  
20196006  
10731132  
2 (bases 1 to 289090)

REFERENCE  
PUBMED  
Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,J.C.

AUTHORS  
Direct Submission

TITLE  
Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA

FEATURES  
source  
Location/Qualifiers  
1. 289090  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/chromosome="x"  
complement(<9373..>16587)  
/gene="EG:BACH48C10.4"  
/note="CG14047"  
/map="2F5-2F5"  
/db\_xref="FLYBASE:Fban0014047"  
/db\_xref="FBgn0040390"  
complement(join(9373..9459,9519..9720,9784..10373,10559..11005,11151..11813,12630..12711,13406..13521,13582..14921,15552..16167,16573..>16587))  
/gene="EG:BACH48C10.4"  
/product="CT33606"  
/db\_xref="FLYBASE:Fban0014047"  
/db\_xref="FBgn0040390"  
complement(join(9373..9459,9519..9720,9784..10373,10559..11005,11151..11813,12630..12711,13406..13521,13582..14921,15552..16167,16573..16587))  
/gene="EG:BACH48C10.4"  
/note="EG:BACH48C10.4 gene product"

CDS

gene  
complement(<18933..>26541)  
/gene="EG:BACH7M4.1"  
/note="CG14045"  
/map="2F5-2F5"  
/db\_xref="FLYBASE:Fban0014045"  
/db\_xref="FBgn0040387"  
complement(join(<18933..20031,20093..20270,20329..21208,21574..22759,24311..24418,25054..25330,25668..26541))  
/gene="EG:BACH7M4.1"  
/product="CT33604"  
/db\_xref="FLYBASE:Fban0014045"  
/db\_xref="FBgn0040387"  
complement(join(18945..20031,20093..20270,20329..21208,21574..22759,24311..24418,25054..25330,25668..26541))  
/gene="EG:BACH7M4.1"  
/note="EG:BACH7M4.1 gene product"  
/codon\_start=1  
/protein\_id="AAE45778.1"  
/db\_xref="GI:7290317"  
/db\_xref="FLYBASE:Fban0014047"  
/db\_xref="FBgn0040390"  
/translation="MDSIIAAGAGAGVVEAGHGHPNVTYITLTFPORTNNSHGHTAGVADTRADRSHGHHGHHGVEDELIDELDKCIIILTKTPKSAHLHANNKQGLDKQOQARLAALARSQENVASADELIFNSDEQERITLTKTRQOLEKTKELNLAKENGGLSAPNFQKRLSGVEELMSASTSDYSTSEAGVATYIKAVDSVGMNPSRSAGEGNDPVITGSRRLHLPDPEVSYSILPNHSGAGVGGSPKPLPPRRYVAMOCPPRRKANOQKATVITETARIRSSPTKSGSGATGSGVSGIKAAACHRCPPEDTHAOLSPDKMEHOTCKLLESPKQTTLTLEKQDEQLSIMLILAOALPAATLCOGERIPKEENSTPIAVVPPDSVLTKTSTHVDNCGSSNGGCTTSTAITYTQKQRAIINIPDSCDESPLEEDPYRMAWSLRFRTWSSLKPLSDRMEDELDDLDSEKNGHDGSPONEDEEDSRALVQNDLGAVSSEGVIVNEGLASGATNRAAGAVSAPMSFEEESRAFIKYLRRMAQAQOHOQTASSTDEQDECTSGVTEGWCTPTSGGNDODOLILNKTSSPTKSLSLDNDDETLEMLLELMAIPMTASTIRGLLRRLREPLFEETESDEKKTQODSDIKGATNNGHLEDSAGSSDSSESPAPPIYPRVYPPRDEPSAEGQENLAMPDQILHSSVRLPNIATSMGPPRLTRLLPTTAIDPPPAVLVASSCEYLIDRPSVPGDLAAGSGITTMNTRSPPLSPAKAAAPRIGMNAVTMTTKTMTTPSRASSTISASDTSSEGRASGRPREVPPPPRRRLITQTDIAQAQKWLGGRRRRORGPMPYRYKDAVANTEYQDLGSGCKTIVINDRDLPAAGTQARLSHIFNSLNGKADPESTIPAEVASELPGCSASTIDGCARTHCNAPRMWRPKPKAGNAAGPQRRSRGGLDLAAMSNDYSGNDPTVNRPRPLSAGLCGTSNPNPNSSSHGSGPLGGCITRAGVEPYGQOOLVYQAGGSAAGVGGSGAGNSTCSESRNGSKGLDEILARRKITYMSQSRCAVRETTSTTTLAAMATEMSMDSEKRDSSWPKRAAAALTRTRSTMDAGAITSTTTTMTTTEAEFLVETVEVSKRLINFRNSETIASTSTSPNPSDGAIGQPSHSQSSLLLOKLPQGOQADQSVEMAQMSPRKNDKPSLNRILMKQITKRRRTNVSQIVAG"

CDS



```
QY 1 TTAGTCTTAGTCTTAGTT 20
Db 1 TTAGTCTTAGTCTTAGTT 20

RESULT 11
AX057376 20 bp DNA linear PAT 17-JAN-2001
LOCUS Sequence 10 from Patent WO0075304.
ACCESSION AX057376
VERSION AX057376.1 GI:12310117
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy,M., Sodoyer,R. and Tranoy,E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 10 14-DEC-2000;
Aventis Pasteur (FR)
FEATURES
source Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="oligonucleotide"

BASE COUNT 3 a 1 c 3 g 13 t
ORIGIN
Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 2.7e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTAGTT 20
Db 1 TTAGTCTTAGTCTTAGTT 20

RESULT 12
AX057379 20 bp DNA linear PAT 17-JAN-2001
LOCUS Sequence 13 from Patent WO0075304.
ACCESSION AX057379
VERSION AX057379.1 GI:12310120
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy,M., Sodoyer,R. and Tranoy,E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 13 14-DEC-2000;
Aventis Pasteur (FR)
FEATURES
source Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="oligonucleotide"

BASE COUNT 3 a 1 c 4 g 12 t
ORIGIN
Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 2.7e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTAGTT 20
Db 1 TTAGTCTTAGTCTTAGTT 20

RESULT 13
AX057381 20 bp DNA linear PAT 17-JAN-2001
LOCUS Sequence 15 from Patent WO0075304.
DEFINITION
```

```
ACCESSION AX057381 GI:12310122
VERSION AX057381.1
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy,M., Sodoyer,R. and Tranoy,E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 15 14-DEC-2000;
Aventis Pasteur (FR)
FEATURES
source Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="oligonucleotide"

BASE COUNT 3 a 1 c 4 g 12 t
ORIGIN
Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 2.7e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTAGTT 20
Db 1 TTAGTCTTAGTCTTAGTT 20

RESULT 14
AC093782 36942 bp DNA linear PRI 01-MAR-2002
LOCUS Homo sapiens BAC clone RP11-20806 from 4, complete sequence.
ACCESSION AC093782 AC009585
VERSION AC093782.3 GI:15920126
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 36942)
Sulston,J.E. and Waterston,R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
JOURNAL MEDLINE
PUBMED 99063792
9847074
REFERENCE 2 (bases 1 to 36942)
AUTHORS Abbott,S., Cotton,M. and Doeber,A.
TITLE The sequence of Homo sapiens BAC clone RP11-20806
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 36942)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (10-SEP-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 36942)
Waterston,R.H.
JOURNAL Direct Submission
Submitted (04-OCT-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 36942)
Waterston,R.
JOURNAL Direct Submission
Submitted (01-MAR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Oct 4, 2001 this sequence version replaced gi:15799641.
COMMENT
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@wustl.wustl.edu
----- Summary Statistics
```



Center project name: H\_NH0208006  
Drafting Center: WIBR

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

**MAPPING INFORMATION:** Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

**SOURCE INFORMATION:**  
The Rpci-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E., Tatematsu, M., Carlsene, J. J. and de Jong, P. J. (1988). An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 5:1-8. The clone name was obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:  
The clone sequenced to the left is RP11-45L9, 2000 bp overlap.  
Actual end of this clone is at base position 36942 of RP11-20806.

Sequence derived from a single plasmid subclone, base position 8833 to 8881.

The sequence of AC009585 has been incorporated into AC093782.

## FEATURES

### Location/Qualifiers

**Source**

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-20806"
/clone_lib="RPCT-11"
275. 349
/rpt_family="MIR"
568. 903
/rpt_family="MALR"
1269. 1635
/rpt_family="MALR"
1830. 2019
/rpt_family="L1"
2633. 2653
/rpt_family="AT_rich"
2714. 3197
/rpt_family="L2"
3209. 3653
/rpt_family="MALR"
3752. 3857
/rpt_family="L2"
4699. 4741
/rpt_family="MALR"
4829. 5153
/rpt_family="L2"
5623. 5927
/rpt_family="L2"
6614. 6933
/rpt_family="L2"

```

repeat_region	6918. .7153	/rpt_family="L2"
misc_feature	7875. .9228	/note="CpC_island (%G=-69.6, o/e=0.80, #CpGs=136) "
misc_feature	8042. .8143	/note="similar to EST AI316961 (MID:g4032228) u124g08.y1"
misc_feature	8064. .8306	/note="similar to EST AI383602 (MID:g4196383) tct46c05.x1"
misc_feature	8068. .8339	/note="similar to EST BF308711 (MID:g11255965) "
misc_feature	8189. .8339	/note="similar to EST AI316961 (MID:g4032228) u124g08.y1"
misc_feature	8245. .8339	/note="similar to EST AI786568 (MID:g5334284) u17e07.y1"
misc_feature	8352. .8719	/note="similar to EST AM027638 (MID:g5886394) wv85g12.x1"
misc_feature	10512. .10628	/note="similar to EST AA913032 (MID:g3052424) ol34a06.s1"
misc_feature	10517. .10628	/note="similar to EST BF308711 (MID:g11255965) "
misc_feature	10517. .10617	/note="similar to EST AI786568 (MID:g5334284) u124g08.y1"
repeat_region	10901. .11008	/rpt_family="ERV1"
repeat_region	10907. .10996	/rpt_family="(TC)n"
repeat_region	10935. .11023	/rpt_family="ERV1"
repeat_region	13451. .13470	/rpt_family="(T)n"
repeat_region	16302. .16399	/rpt_family="L1"
repeat_region	16400. .16699	/rpt_family="Alu"
repeat_region	16700. .16940	/rpt_family="L1"
repeat_region	17303. .17650	/rpt_family="Alu"
repeat_region	18082. .18352	/rpt_family="L1"
misc_feature	18614. .18895	/note="similar to EST BF762649 (MID:g12110549) "
misc_feature	18942. .19369	/note="similar to EST BG977308 (MID:g14380043) "
misc_feature	19505. .19923	/note="similar to EST AI936059 (MID:g5674929) wo47g12.x1"
repeat_region	21186. .21476	/rpt_family="Alu"
repeat_region	23528. .23925	/note="similar to EST BF328962 (MID:g11299710) "
misc_feature	24479. .24854	/note="similar to EST AI907471 (MID:g6497901) "
misc_feature	24804. .24854	/note="similar to EST AI786568 (MID:g5334284) u17e07.y1"
repeat_region	25482. .25781	/rpt_family="Alu"
repeat_region	25847. .25868	/rpt_family="AT_rich"
repeat_region	26045. .26132	/rpt_family="MIR"
repeat_region	26578. .26666	/rpt_family="L2"
repeat_region	26956. .26989	/rpt_family="(TC)n"
repeat_region	26990. .27097	/rpt_family="(TA)n"
misc_feature	27529. .27954	/note="similar to EST AM027811 (MID:g5886567) wv60e01.x1"
repeat_region	28028. .28207	

repeat\_region /rpt\_family="MIR"  
28306..28419 /rpt\_family="L2"  
repeat\_region 28838..29136 /rpt\_family="Alu"  
repeat\_region 28997..29018 /rpt\_family="Alu"  
repeat\_region 29200..29460 /rpt\_family="Ar\_rich"  
/rpt\_family="L2"  
misc\_feature 29573..29782 /note="similar to EST BF758076 (NID:g12105963)"

Query Match 92.0%; Score 18.4; DB 8; Length 36942;  
Best Local Similarity 95.0%; Pred. No. 64;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTCTTAGTT 20  
|||||  
Db 2953 TTAGTCTTAGTCTTAGTT 2972

RESULT 15  
AC092116

LOCUS AC092116 61372 bp DNA linear PRI 05-MAR-2002  
DEFINITION Homo sapiens chromosome 16 clone CTD-2133M23, complete sequence.

ACCESSION AC092116

VERSION AC092116.2 GI:19111661

KEYWORDS HTG.

SOURCE Homo sapiens.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 61372)

AUTHORS DOE Joint Genome Institute.

TITLE Sequencing of Human Chromosome 16

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 61372)

AUTHORS DOE Joint Genome Institute.

TITLE Direct Submission

JOURNAL Submitted (22-JUN-2001) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCE 3 (bases 1 to 61372)

AUTHORS DOE Joint Genome Institute.

TITLE Direct Submission

JOURNAL Submitted (05-MAR-2002) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

On Mar 5, 2002 this sequence version replaced gi:14522962.

COMMENT

Sequence Quality Assessment:  
This entry has been annotated with sequence quality

estimates computed by the Phrap assembly program.

All manually edited bases have been reduced to quality zero.

Quality levels above 40 are expected to have less than

1 error in 10,000 bp.

Base-by-base quality values are not generally visible from the

GenBank flat file format but are available as part

of this entry's ASN.1 file.

-----  
Sequence Quality Assessment:

This entry has been annotated with sequence quality

estimates computed by the Phrap assembly program.

All manually edited bases have been reduced to quality zero.

Quality levels above 40 are expected to have less than

1 error in 10,000 bp.

Base-by-base quality values are not generally visible from the

GenBank flat file format but are available as part

of this entry's ASN.1 file.

-----

FEATURES

Source

Location/Qualifiers

1..61372

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="16"

/clone="CTD-2133M23"

BASE COUNT 19233 a 11657 c 11252 g 19230 t  
ORIGIN

Query Match

Best Local Similarity 92.0%; Score 18.4; DB 8; Length 61372;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTCTTAGTT 20  
|||||  
Db 27299 TTAACTCTAGTCTTAGTT 27318

Search completed: April 6, 2003, 11:56:41  
Job time : 418 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 11:21:32 ; Search time 151 Seconds  
(without alignments)  
298.278 Million cell updates/sec

Title: US-09-980-265-11

Perfect score: 20

Sequence: 1 ttacttcttagttcttagtt 20

Scoring table:

IDENTITY-NUC  
Gapop 10.0 , Gapept 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

N.GeneSeq\_101002:\*

1: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1980.DAT:\*

2: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1981.DAT:\*

3: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1982.DAT:\*

4: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1983.DAT:\*

5: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1984.DAT:\*

6: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1985.DAT:\*

7: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1986.DAT:\*

8: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1987.DAT:\*

9: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1988.DAT:\*

10: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1989.DAT:\*

11: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1990.DAT:\*

12: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1991.DAT:\*

13: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1992.DAT:\*

14: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1993.DAT:\*

15: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1994.DAT:\*

16: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1995.DAT:\*

17: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1996.DAT:\*

18: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1997.DAT:\*

19: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1998.DAT:\*

20: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1999.DAT:\*

21: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA2000.DAT:\*

22: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA2001A.DAT:\*

23: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA2001B.DAT:\*

24: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	22	AA28882
2	19	95.0	20	22	AA28871
3	18.4	92.0	20	22	AA28874
4	18.4	92.0	20	22	AA28878
5	18.4	92.0	20	22	AA28880
6	18.4	92.0	20	22	AA28881
7	18.4	92.0	20	22	AA28883
8	18.4	92.0	20	22	AA28886
9	18.4	92.0	4741	23	ABL27984

C	10	18.4	92.0	13798	23	ABL08500
C	11	18.4	92.0	31034	23	ABL13678
C	12	18	90.0	32183	22	AA199267
C	13	18	90.0	32183	22	AA199266
C	14	18	90.0	32192	22	AA199266
C	15	18	90.0	32192	22	AA199266
C	16	16.8	84.0	20	22	AA28872
C	17	16.8	84.0	20	22	AA28873
C	18	16.8	84.0	20	22	AA28875
C	19	16.8	84.0	20	22	AA28876
C	20	16.8	84.0	20	22	AA28877
C	21	16.8	84.0	20	22	AA28879
C	22	16.8	84.0	20	22	AA28884
C	23	16.8	84.0	20	22	AA28885
C	24	16.8	84.0	20	22	AA28887
C	25	16.8	84.0	371	22	AA65716
C	26	16.8	84.0	377	22	AA65717
C	27	16.8	84.0	414	23	ABV47200
C	28	16.8	84.0	490	23	ABV17406
C	29	16.8	84.0	572	22	AAK08181
C	30	16.8	84.0	572	22	AAK34060
C	31	16.8	84.0	572	22	AAI39783
C	32	16.8	84.0	572	24	AB508840
C	33	16.8	84.0	749	24	ABQ21728
C	34	16.8	84.0	749	24	ABQ21729
C	35	16.8	84.0	1069	21	AAK46432
C	36	16.8	84.0	1071	21	AAK39993
C	37	16.8	84.0	1240	21	AAK45636
C	38	16.8	84.0	1242	21	AAK40226
C	39	16.8	84.0	4866	23	ABL04084
C	40	16.8	84.0	4920	18	AAV74337
C	41	16.8	84.0	5453	24	ABL70450
C	42	16.8	84.0	5453	24	AAK56143
C	43	16.8	84.0	5453	24	ABK31481
C	44	16.8	84.0	7238	24	ABL32448
C	45	16.8	84.0	7353	24	ABL32072

#### ALIGNMENTS

RESULT 1	AA28882	standard: DNA; 20 BP.
ID	AA28882	
AC	AA28882	
DT	09-MAY-2001	(first entry)
DE	Immunostimulatory oligonucleotide #11 as vaccine adjuvant.	
KW	Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;	
KW	prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;	
KW	cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;	
KW	phosphorothioate; ss.	
OS	Synthetic.	
FH	Key	Location/Qualifiers
FT	modified_base	1..20
FT		/*tag= a
FT		/note= "contain phosphorothioate internucleotide linkages"
PN	WO200075304-A1.	
PD	14-DEC-2000.	
PP	08-JUN-2000; 2000WO-FR01566.	
PR	08-JUN-1999; 99FR-0007457.	
PR	06-AUG-1999; 99FR-0010378.	
XX		

Drosophila melanog  
Drosophila melanog  
Human excretory re  
Human kidney relat  
Human excretory re  
Human kidney relat  
Immunostimulatory  
Immunostimulatory  
Immunostimulatory  
Immunostimulatory  
Immunostimulatory  
Immunostimulatory  
Immunostimulatory  
Human immune/haema  
Human immune/haema  
Human prostate exp  
Human prostate exp  
Human brain expres  
Human bone marrow  
Probe #8469 used t  
Human genome-deriv  
Oligonucleotide fo  
Oligonucleotide fo  
Arabidopsis thalia  
Arabidopsis thalia  
Arabidopsis thalia  
Drosophila melanog  
Staphylococcus aur  
Chemically treated  
Human gene regulat  
Signal transductio  
Human immune syste  
Human immune syste

PA (AVET ) AVENTIS PASTEUR.  
 PI Bachy M, Sodayer R, Tranoy E;  
 XX WPI; 2001-041317/05.  
 DR  
 XX  
 PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in  
 PT vaccines for human use, induce lymphocyte proliferation and cytokine  
 PT secretion -  
 XX  
 PS Example 4; Page 13; 30pp; French.  
 XX  
 CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an  
 CC immunostimulatory oligonucleotide of the invention which contains at  
 CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The  
 CC oligonucleotides do not contain any CG dinucleotides in which C are  
 CC unethylated. The oligonucleotides are used as human immunostimulants  
 CC and as adjuvants in therapeutic and prophylactic vaccines for human use.  
 CC They induce proliferation of human lymphocytes, induce secretion of  
 CC cytokines, especially interleukin-10 or interferon-gamma and increase  
 CC expression of the CD86 activation marker or the CD25 cytokine receptor  
 CC on human B lymphocytes.  
 CC  
 XX  
 SO Sequence 20 BP; 3 A; 2 C; 3 G; 12 T; 0 other;  
 Query Match 100.0%; Score 20; DB 22; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 8.8;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 TTAGTCTTAGTCTTAGTT 20  
 1 TTAGTCTTAGTCTTAGTT 20  
 Db  
 RESULT 2  
 AAF28871  
 ID AAF28871 standard; DNA: 20 BP.  
 XX  
 AC AAF28871;  
 XX  
 DT 09-MAY-2001 (first entry)  
 XX  
 DE Generic immunostimulatory oligonucleotide as vaccine adjuvant.  
 XX  
 KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;  
 KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;  
 KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor; ss.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200075304-A1.  
 PD 14-DEC-2000.  
 XX  
 PE 08-JUN-2000; 2000MO-FR01566.  
 XX  
 PR 08-JUN-1999; 99FR-0007457.  
 PR 06-AUG-1999; 99FR-0010378.  
 XX  
 PA (AVET ) AVENTIS PASTEUR.  
 XX  
 PI Bachy M, Sodayer R, Tranoy E;  
 XX  
 DR WPI; 2001-041317/05.  
 XX  
 PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in  
 PT vaccines for human use, induce lymphocyte proliferation and cytokine  
 PT secretion -  
 XX  
 PS Claim 8; Page 17; 30pp; French.  
 CC This sequence represents a generic example of an immunostimulatory  
 CC oligonucleotide of the invention which contains at least one sequence

CC 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The oligonucleotides  
 CC do not contain any CG dinucleotides in which C are unethylated.  
 CC The oligonucleotides are used as human immunostimulants and as adjuvants  
 CC in therapeutic and prophylactic vaccines for human use. They induce  
 CC proliferation of human lymphocytes, induce secretion of cytokines,  
 CC especially interleukin-10 or interferon-gamma and increase expression  
 CC of the CD86 activation marker or the CD25 cytokine receptor on human B  
 CC lymphocytes.  
 CC  
 XX  
 SO Sequence 20 BP; 3 A; 1 C; 3 G; 12 T; 1 other;  
 Query Match 95.0%; Score 19; DB 22; Length 20;  
 Best Local Similarity 95.0%; Pred. No. 24;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 TTAGTCTTAGTCTTAGTT 20  
 1 TTAGTCTTAGTCTTAGTT 20  
 Db 1 TTAGTCTTAGTCTTAGTT 20  
 RESULT 3  
 AAF28874  
 ID AAF28874 standard; DNA: 20 BP.  
 XX  
 AC AAF28874;  
 XX  
 DT 09-MAY-2001 (first entry)  
 XX  
 DE Immunostimulatory oligonucleotide #3 as vaccine adjuvant.  
 XX  
 KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;  
 KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;  
 KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;  
 KW phosphorothioate; ss.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200075304-A1.  
 PD 14-DEC-2000.  
 XX  
 PE 08-JUN-2000; 2000MO-FR01566.  
 XX  
 PR 08-JUN-1999; 99FR-0007457.  
 PR 06-AUG-1999; 99FR-0010378.  
 XX  
 PA (AVET ) AVENTIS PASTEUR.  
 XX  
 PI Bachy M, Sodayer R, Tranoy E;  
 XX  
 DR WPI; 2001-041317/05.  
 XX  
 PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in  
 PT vaccines for human use, induce lymphocyte proliferation and cytokine  
 PT secretion -  
 XX  
 PS Example 4; Page 13; 30pp; French.  
 CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an  
 CC immunostimulatory oligonucleotide of the invention which contains at  
 CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The  
 CC oligonucleotides do not contain any CG dinucleotides in which C are  
 CC unethylated. The oligonucleotides are used as human immunostimulants  
 CC and as adjuvants in therapeutic and prophylactic vaccines for human use.  
 CC They induce proliferation of human lymphocytes, induce secretion of  
 CC cytokines, especially interleukin-10 or interferon-gamma and increase  
 CC expression of the CD86 activation marker or the CD25 cytokine receptor

CC on human B lymphocytes.

XX Sequence 20 BP; 4 A; 1 C; 3 G; 12 T; 0 other:

Query Match

Best Local Similarity 92.0%; Score 18.4; DB 22; Length 20;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTCTTAGTT 20  
||||| |||||||||  
Db 1 TTAGTTATTAGTTCTTAGTT 20

RESULT 4

AAF28878 ID AAF28878 standard; DNA; 20 BP.

AC AAF28878;

DT 09-MAY-2001 (first entry)

XX Immunostimulatory oligonucleotide #7 as vaccine adjuvant.

KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;  
KM prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;  
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;  
KW phosphorothioate; ss.

XX Synthetic.

OS Key Location/Qualifiers

FH modified\_base 1..20

FT /tag= a  
FT /note= "contain phosphorothioate internucleotide linkages"

FT WO200075304-A1.

PD 14-DEC-2000.

PE 08-JUN-2000; 2000WO-FR01566.

PR 08-JUN-1999; 99FR-0007457.

PR 06-AUG-1999; 99FR-0010378.

XX (AVET ) AVENTIS PASTEUR.

PA Bachy M, Sodyer R, Tranoy E;

PI WPI; 2001-041317/05.

PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in  
PT vaccines for human use, induce lymphocyte proliferation and cytokine  
PT secretion -

PS Example 4; Page 13; 30pp; French.

CC Oligonucleotides AAF28872-AAF2887 represent specific examples of an  
CC immunostimulatory oligonucleotide of the invention which contains at  
CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The  
CC oligonucleotides do not contain any CG dinucleotides in which C are  
CC unmethylated. The oligonucleotides are used as human immunostimulants  
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.  
CC They induce proliferation of human lymphocytes, induce secretion of  
CC cytokines, especially interleukin-10 or interferon-gamma and increase  
CC expression of the CD86 activation marker or the CD25 cytokine receptor  
CC on human B lymphocytes.

SO Sequence 20 BP; 3 A; 1 C; 3 G; 13 T; 0 other:

Query Match

Best Local Similarity 92.0%; Score 18.4; DB 22; Length 20;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTCTTAGTT 20  
||||| |||||||||  
Db 1 TTAGTTATTAGTTCTTAGTT 20

RESULT 5

AAF28880 ID AAF28880 standard; DNA; 20 BP.

AC AAF28880;

DT 09-MAY-2001 (first entry)

XX Immunostimulatory oligonucleotide #9 as vaccine adjuvant.

KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;  
KM prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;  
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;  
KW phosphorothioate; ss.

XX Synthetic.

OS Key Location/Qualifiers

FH modified\_base 1..20

FT /tag= a  
FT /note= "contain phosphorothioate internucleotide linkages"

FT WO200075304-A1.

PD 14-DEC-2000.

PE 08-JUN-2000; 2000WO-FR01566.

PR 08-JUN-1999; 99FR-0007457.

PR 06-AUG-1999; 99FR-0010378.

XX (AVET ) AVENTIS PASTEUR.

PA Bachy M, Sodyer R, Tranoy E;

PI WPI; 2001-041317/05.

PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in  
PT vaccines for human use, induce lymphocyte proliferation and cytokine  
PT secretion -

PS Example 4; Page 13; 30pp; French.

CC Oligonucleotides AAF28872-AAF2887 represent specific examples of an  
CC immunostimulatory oligonucleotide of the invention which contains at  
CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The  
CC oligonucleotides do not contain any CG dinucleotides in which C are  
CC unmethylated. The oligonucleotides are used as human immunostimulants  
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.  
CC They induce proliferation of human lymphocytes, induce secretion of  
CC cytokines, especially interleukin-10 or interferon-gamma and increase  
CC expression of the CD86 activation marker or the CD25 cytokine receptor  
CC on human B lymphocytes.

SO Sequence 20 BP; 4 A; 1 C; 3 G; 12 T; 0 other:

Query Match

Best Local Similarity 92.0%; Score 18.4; DB 22; Length 20;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTCTTAGTT 20  
||||| |||||||||  
Db 1 TTAGTTATTAGTTATTAGTT 20

RESULT 6

```
AAF28881
ID AAF28881 standard; DNA: 20 BP.
XX
XX AAF28881;
AC
XX
DT 09-MAY-2001 (first entry)
DE Immunostimulatory oligonucleotide #10 as vaccine adjuvant.
XX
XX Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KM prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
XX phosphorothioate; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1..20
FT /tag= a
FT /note= "contain phosphorothioate internucleotide
FT linkages"
XX
XX WO200075304-A1.
XX
XX 14-DEC-2000.
XX
XX 08-JUN-2000; 2000WO-FR01566.
XX
XX 08-JUN-1999; 99FR-0007457.
XX 06-AUG-1999; 99FR-0010378.
XX
XX (AVET ) AVENTIS PASTEUR.
XX
XX Bachy M, Sodoyer R, Tranoy E;
XX
XX WPI; 2001-041317/05.
XX
XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
PT vaccines for human use, induce lymphocyte proliferation and cytokine
PT secretion -
XX
XX Example 4; Page 13; 30pp; French.
XX
XX Oligonucleotides AAF28872-AAF28887 represent specific examples of an
CC immunostimulatory oligonucleotide of the invention which contains at
CC least one sequence 5'-TTNN2YT-3' where N1 and N2 are A, T, C or G. The
CC oligonucleotides do not contain any CG dinucleotides in which C are
CC unethylated. The oligonucleotides are used as human immunostimulants
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
CC They induce proliferation of human lymphocytes, induce secretion of
CC cytokines, especially interleukin-10 or interferon-gamma and increase
CC expression of the CD86 activation marker or the CD25 cytokine receptor
CC on human B lymphocytes.
XX
XX Sequence 20 BP; 3 A; 1 C; 3 G; 13 T; 0 other;
SQ
Query Match 92.0%; Score 18.4; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 43;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTAGTCTCTAGTCTTAGT 20
DB 1 TTAGTCTCTAGTCTTAGT 20
RESULT 7
AAF28883
ID AAF28883 standard; DNA: 20 BP.
XX
XX AAF28883;
AC
XX
XX 09-MAY-2001 (first entry)
DT
```

```
DE Immunostimulatory oligonucleotide #12 as vaccine adjuvant.
XX
XX Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KM prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
XX phosphorothioate; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1..20
FT /tag= a
FT /note= "contain phosphorothioate internucleotide
FT linkages"
XX
XX WO200075304-A1.
XX
XX 14-DEC-2000.
XX
XX 08-JUN-2000; 2000WO-FR01566.
XX
XX 08-JUN-1999; 99FR-0007457.
XX 06-AUG-1999; 99FR-0010378.
XX
XX (AVET ) AVENTIS PASTEUR.
XX
XX Bachy M, Sodoyer R, Tranoy E;
XX
XX WPI; 2001-041317/05.
XX
XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
PT vaccines for human use, induce lymphocyte proliferation and cytokine
PT secretion -
XX
XX Example 4; Page 13; 30pp; French.
XX
XX Oligonucleotides AAF28872-AAF28887 represent specific examples of an
CC immunostimulatory oligonucleotide of the invention which contains at
CC least one sequence 5'-TTNN2YT-3' where N1 and N2 are A, T, C or G. The
CC oligonucleotides do not contain any CG dinucleotides in which C are
CC unethylated. The oligonucleotides are used as human immunostimulants
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
CC They induce proliferation of human lymphocytes, induce secretion of
CC cytokines, especially interleukin-10 or interferon-gamma and increase
CC expression of the CD86 activation marker or the CD25 cytokine receptor
CC on human B lymphocytes.
XX
XX Sequence 20 BP; 3 A; 1 C; 4 G; 12 T; 0 other;
SQ
Query Match 92.0%; Score 18.4; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 43;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTAGTCTCTAGTCTTAGT 20
DB 1 TTAGTCTCTAGTCTTAGT 20
RESULT 8
AAF28886
ID AAF28886 standard; DNA: 20 BP.
XX
XX AAF28886;
AC
XX
XX 09-MAY-2001 (first entry)
DT
DE Immunostimulatory oligonucleotide #15 as vaccine adjuvant.
XX
XX Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KM prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
XX phosphorothioate; ss.
XX
```

OS	Synthetic.
XX	
FH	Key
FT	Location/Qualifiers
FT	modified_base
FT	1..20
FT	/tag= a
FT	/note= "contain phosphorothioate internucleotide linkages"
XX	
PN	WO20075304-A1.
XX	
PD	14-DEC-2000.
XX	
PE	08-JUN-2000; 2000WO-FR01566.
XX	
PR	08-JUN-1999; 99FR-0007457.
PR	06-AUG-1999; 99FR-0010378.
XX	
PA	(AVET ) AVENTIS PASTEUR.
XX	
PI	Bachy M, Sodoyer R, Trannoy E;
XX	
DR	WPI; 2001-041317/05.
PT	New immunostimulatory oligonucleotides, useful e.g. as adjuvants in vaccines for human use, induce lymphocyte proliferation and cytokine secretion -
PT	
XX	
PS	Example 4; Page 13; 30pp; French.
XX	
CC	Oligonucleotides AAF28672-AAF28887 represent specific examples of an immunostimulatory oligonucleotide of the invention which contains at least one sequence 5'-TTNIN2T-3' where N1 and N2 are A, T, C or G. The oligonucleotides do not contain any CG dinucleotides in which C are unmethylated. The oligonucleotides are used as human immunostimulants and as adjuvants in therapeutic and prophylactic vaccines for human use. They induce proliferation of human lymphocytes, induce secretion of cytokines, especially interleukin-10 or interferon-gamma and increase expression of the CD86 activation marker or the CD25 cytokine receptor on human B lymphocytes.
CC	
CC	Sequence 20 BP; 3 A; 1 C; 4 G; 12 T; 0 other;
XX	
SO	
Query Match	92.0%; Score 18.4; DB 22; Length 20;
Best Local Similarity	95.0%; Pred. NO. 43;
Matches 19; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
QY	1 TTACTTCTTAGTCTTACGT 20             1 TTAGTTGTTAGTCTTACGT 20
Db	
RESULT 9	
ABL27984/C	
ID	ABL27984 standard; DNA; 4741 BP.
XX	
AC	ABL27984:
XX	
DT	26-MAR-2002 (first entry)
DE	Drosophila melanogaster genomic polynucleotide SEQ ID NO 35425.
XX	
KM	Drosophila; developmental biology; cell signalling; insecticide;
KW	pharmaceutical; gene; ds.
XX	
OS	Drosophila melanogaster.
XX	
PN	WO200171042-A2.
XX	
PD	27-SEP-2001.
XX	
PF	23-MAR-2001; 2001WO-US09231.
XX	
PR	23-MAR-2000; 2000US-191637P.

```

PR      11-JUL-2000; 2000US-0614150.
PA      (PEKE ) PE CORP NY.
XX
XX      Venter JC, Adams M, Li PWD, Myers EW;
PI
DR      WPI; 2001-656860/75.
XX
XX      New isolated nucleic acid detection reagent for detecting 1000 or more
PT      genes from Drosophila and for elucidating cell signalling and cell-cell
PT      interactions -
XX
XX
PS      Claim 1; SEQ ID NO 35425; 21pp + Sequence Listing; English.
CC
CC      The invention relates to an isolated nucleic acid detection reagent
CC      capable of detecting 1000 or more genes from Drosophila. The invention is
CC      useful in developmental biology and in elucidating cell signalling and
CC      cell-cell interactions in higher eukaryotes for the development of
CC      insecticides, therapeutics and pharmaceutical drugs. The invention
CC      discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC      sequences (ABL01840-ABL16175) and the encoded proteins
CC
CC      The sequence data for this patent did not form part of the printed
CC      specification, but was obtained in electronic format directly from WIPO
CC      at ftp.wipo.int/pub/published_pct_sequences.
XX
XX      Sequence 4741 BP; 1155 A; 925 C; 1234 G; 1427 T; 0 other;
SQ
XX
XX      Query Match          92.0%; Score 18.4; DB 23; Length 4741;
XX      Best Local Similarity 95.0%; Pred. No. 50;
XX      Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX      1 TTAGTCTTACTTCTTACTT 20
XX      I | | | | | | | | | | | | | |
XX      Db 356 TAACTTCTTACTTCTTACTT 337
XX
XX
XX      RESULT 10
XX      ABL08500/c
XX      ID ABL08500 standard; cDNA; 13798 BP.
XX
XX      ABL08500;
XX
XX      26-MAR-2002 (first entry)
XX
XX      Drosophila melanogaster expressed polynucleotide SEQ ID NO 19982.
XX
XX      Drosophila; developmental biology; cell signalling; insecticide;
XX      pharmaceutical; gene; ss.
XX
XX      Drosophila melanogaster.
XX
XX      WO200171042-A2.
XX
XX      27-SEP-2001.
XX
XX      23-MAR-2001; 2001WO-US09231.
XX
XX      23-MAR-2000; 2000US-191637P.
XX      PR 11-JUL-2000; 2000US-0614150.
XX
XX      (PEKE ) PE CORP NY.
XX
XX      Venter JC, Adams M, Li PWD, Myers EW;
XX
XX      WPI; 2001-656860/75.
XX      DR P-PsDE; ABB64397.
XX
XX      New isolated nucleic acid detection reagent for detecting 1000 or more
XX      genes from Drosophila and for elucidating cell signalling and cell-cell
XX      interactions -
XX
XX      Claim 1; SEQ ID NO 19982; 21pp + Sequence Listing; English.

```

XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA  
CC sequences (AB101840-AB116175) and the encoded proteins  
CC (ABBS7737-ABBS72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).  
SQ Sequence 13798 BP; 4299 A; 2926 C; 2679 G; 3894 T; 0 other;  
Query Match 92.0%; Score 18.4; DB 23; Length 13798;  
Best Local Similarity 95.0%; Pred. No. 51;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 TTACTTCTTACTCTTACTT 20  
Db 9698 TTACTTCTTACTCTTACTT 9679  
||||| |||||||  
RESULT 11  
AB13678/c  
ID AB13678 standard; cDNA; 31034 BP.  
AC ABL13678;  
XX  
XX 26-MAR-2002 (first entry)  
DE *Drosophila melanogaster* expressed polynucleotide SEQ ID NO 35516.  
XX  
XX *Drosophila*: developmental biology; cell signalling; insecticides;  
KM pharmaceutical; gene; ss.  
XX  
XX *Drosophila melanogaster*.  
PN WO200171042-A2.  
XX  
XX 27-SEP-2001.  
PD  
XX  
XX 23-MAR-2001; 2001WO-US09231.  
PE  
XX  
XX 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX  
XX (PEKE ) PE CORP NY.  
PA  
XX Venter JC, Adams M, Li PWD, Myers EW;  
PI WPI; 2001-656860/75.  
XX  
XX P-PSDB; ABB69575.  
DR  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from *Drosophila* and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
XX  
PS Claim 1; SEQ ID NO 35516; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA  
CC sequences (AB101840-AB116175) and the encoded proteins  
CC (ABBS7737-ABBS72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).  
XX

SQ Sequence 31034 BP; 9528 A; 5898 C; 5991 G; 9617 T; 0 other;  
Query Match 92.0%; Score 18.4; DB 23; Length 31034;  
Best Local Similarity 95.0%; Pred. No. 52;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 TTACTTCTTACTCTTACTT 20  
Db 28137 TTACTTCTTACTCTTACTT 28118  
||||| |||||||  
RESULT 12  
AA199267/c  
ID AA199267 standard; DNA; 32183 BP.  
XX  
XX AA199267;  
AC  
XX  
XX 07-JAN-2002 (first entry)  
DE Human excretory related polynucleotide SEQ ID NO 1031.  
XX  
XX Human; neurotropic; neuroprotective; cytosolic; dermatological; virocidic;  
KM immunosuppressive; anti-infectious; anti-HIV; antibacterial; vulnereary;  
KM antiparkinsonian; antispasmodic; antianemic; antiarthritic; cancer;  
KM antitumour; hepatotropic; cerebroprotective; antiinflammatory;  
KM antiparasitic; antidiabetic; antitumor; anticonvulsant; antifungal;  
KM antiparasitic; cardiac; immune disorder; cardiovascular disorder;  
KM neurological disease; infection; nephrotoxic; gene therapy; vaccine;  
XX excretory system; ds.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200155313-A2.  
PN  
XX  
XX 02-AUG-2001.  
PD  
XX  
XX 17-JAN-2001; 2001WO-US01323.  
PE  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 14-AUG-2000; 2000US-0226279.  
PR 18-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR



PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0233397.  
PR 14-SEP-2000; 2000US-0233398.  
PR 14-SEP-2000; 2000US-0233399.  
PR 14-SEP-2000; 2000US-0233400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239335.  
PR 13-OCT-2000; 2000US-0239337.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-024617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246539.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.

PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 11-DEC-2000; 2000US-0251990.  
PR 05-JAN-2001; 2001US-0259678.  
(HUMA-) HUMAN GENOME SCI INC.  
Rosen CA, Barash SC, Ruben SM;  
WPI: 2001-465569/50.  
Isolated nucleic acid molecule encoding excretory system antigen is used in preventing, treating or ameliorating a medical condition -  
Example 2; SEQ ID NO 1031; 574pp + Sequence Listing; English.  
XX  
CC The invention relates to novel excretory system related human  
CC polynucleotides (AA198567-AA199503) and the encoded proteins  
CC (AAAG9594-AAAG9913) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy, especially  
CC disorders related to the excretory system. The genes are isolated  
CC from a range of human tissues disclosed in the specification. The  
CC nucleic acids, proteins, antibodies and (ant)agonists are useful in  
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;  
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 32183 BP; 9177 A; 6710 C; 6857 G; 9439 T; 0 other;  
Query Match 90.0%; Score 18; DB 22; Length 32183;  
Best Local Similarity 100.0%; Pred. No. 78;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTAGTCTTAGTCTTAG 18  
|||||  
Db 17046 TTAGTCTTAGTCTTAG 17029  
RESULT 13

AA163617/c  
ID AA163617 standard; DNA: 32183 BP.  
XX  
AC AA163617;  
XX  
DT 22-OCT-2001 (first entry)  
XX  
DE Human kidney related polynucleotide SEQ ID NO 932.  
XX  
KW Human: kidney antigen; immunosuppressive; antirheptic; antirheumatic;  
KW antiproliferative; cytoskeletal; cardant; vasotropic; cerebroprotective;  
KW neurotropic; neuroprotective; antibacterial; virucide; fungicide;  
KW ophthalmological; antidiabetic; hepatotropic; antidiabetic;  
KW antiinflammatory; antitumor; antiviral; anticonvulsant; antiparasitic;  
KW gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; ds.  
XX  
OS Homo sapiens.  
XX  
PN W0200155323-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001MO-US01343.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 23-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231245.  
PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.

XX	17-NOV-2000;	2000US-0249299.		
XX	17-NOV-2000;	2000US-0249300.		
PR	01-DEC-2000;	2000US-0250160.		
PR	01-DEC-2000;	2000US-0250391.		
PR	05-DEC-2000;	2000US-0251030.		
PR	05-DEC-2000;	2000US-0251988.		
PR	05-DEC-2000;	2000US-0256719.		
PR	06-DEC-2000;	2000US-0251479.		
PR	08-DEC-2000;	2000US-0251856.		
PR	08-DEC-2000;	2000US-0251868.		
PR	08-DEC-2000;	2000US-0251869.		
PR	08-DEC-2000;	2000US-0251989.		
PR	08-DEC-2000;	2000US-0251990.		
PR	11-DEC-2000;	2000US-0254097.		
PR	05-JAN-2001;	2001US-0259678.		
PA	(HUMA-) HUMAN GENOME SCI INC.			
XX				
PI	Rosen CA, Barash SC, Ruben SM;			
DR	WPI: 2001-488784/53.			
XX				
PT	New isolated nucleic acids and polypeptides, useful for diagnosing, treating and/or preventing human diseases and disorders -			
XX				
PS	Disclosure: SEQ ID NO 932; 564bp + Sequence Listing; English.			
XX				
CC	The invention relates to novel kidney related polynucleotides (AA162971-AA163793) and the encoded polypeptides (AAM2417-AAM2691) collectively known as kidney antigens and the use of such kidney antigens for detecting disorders of the kidney, especially kidney cancer and kidney cancer metastases. The polynucleotides and proteins are also useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer, and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d) wound healing; (e) neurological diseases such as cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.			
CC	Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct-sequences.			
XX				
SQ	Sequence 32183 BP; 9177 A; 6710 C; 6857 G; 9439 T; 0 other;			
Query Match				
Best Local Similarity 90.0%; Score 18; DB 22; Length 32183;				
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1 TTAGTCTTAGTCTTAG 18 			
Db	17046 TTAGTCTTAGTCTTAG 17029			
RESULT 14				
ID	AA19266/c			
XX	AA19266 standard; DNA: 32192 BP.			
AC	AA19266;			
XX				
DT	07-JAN-2002 (first entry)			
XX				
DE	Human excretory related polynucleotide SEQ ID NO 1030.			
XX				
XX	Human; nootropic; neuroprotective; cyostatic; dermatological; virucide; immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;			

PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0242223.  
 PR 21-SEP-2000; 2000US-0242274.  
 PR 25-SEP-2000; 2000US-0243997.  
 PR 25-SEP-2000; 2000US-0243998.  
 PR 26-SEP-2000; 2000US-0235464.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236359.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.

PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX WPI, 2001-465569/50.  
 DR  
 XX Isolated nucleic acid molecule encoding excretory system antigen is  
 PT used in preventing, treating or ameliorating a medical condition  
 PS Example 2; SEQ ID NO 1030; 574pp + Sequence Listing; English.  
 XX  
 CC The invention relates to novel excretory system related human  
 CC polynucleotides (AA198567-AA199503) and the encoded proteins  
 CC (AAM99594-AAM99913) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy, especially  
 CC disorders related to the excretory system. The genes are isolated  
 CC from a range of human tissues disclosed in the specification. The  
 CC nucleic acids, proteins, antibodies and (ant)agonists are useful in  
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
 CC hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;  
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
 CC and parasitic infections.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 32192 BP; 8946 A; 6809 C; 6885 G; 9552 T; 0 other;  
 OY 1 TTAGTCTTAGTCTTAG 18  
 DB 24031 TTAGTCTTAGTCTTAG 24014  
 AC  
 XX  
 AC AA163616;  
 XX  
 XX  
 DT 22-OCT-2001 (first entry)  
 XX  
 DE Human kidney related polynucleotide SEQ ID NO 931.  
 XX  
 KW Human; kidney antigen; immunosuppressive; antiarthritis; antirheumatic;  
 KW antiproliferative; cytostatic; cardiant; vasotrophic; cerebroprotective;  
 KW nootropic; neuroprotective; antibacterial; virucide; fungicide;  
 KW ophthalmological; antiallergic; hepatotropic; antidiabetic;  
 KW antiinflammatory; antituber; vulnerable; anticonvulsant; antiparasitic;  
 KW gene therapy; cancer; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; ds.  
 KW  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200155323-A2.  
 XX  
 PD 02-AUG-2001.  
 XX

PF 17-JAN-2001; 2001WO-US01343.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 21-SEP-2000; 2000US-0234065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.

PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239335.  
PR 13-OCT-2000; 2000US-0239337.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0248207.  
PR 17-NOV-2000; 2000US-0248208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 03-JAN-2001; 2001US-0259678.  
  
(HUMA-) HUMAN GENOME SCI INC.  
XX  
PA  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX WPI: 2001-488784/53.  
XX  
PT New isolated nucleic acids and polypeptides, useful for diagnosing,

PT treating and/or preventing human diseases and disorders -  
 XX  
 PS Disclosure; SEQ ID NO 931; 564bp + Sequence Listing; English.  
 XX  
 CC The invention relates to novel kidney related polynucleotides  
 CC (AA162971-AA163793) and the encoded polypeptides (AAM42417-AAM42691)  
 CC collectively known as kidney antigens and the use of such kidney antigens  
 CC for detecting disorders of the kidney, especially kidney cancer and  
 CC kidney cancer metastases. The polynucleotides and proteins are also  
 CC useful for preventing, treating or ameliorating medical conditions  
 CC e.g. by protein or gene therapy. The genes are isolated from a range  
 CC of human tissues disclosed in the specification. The nucleic acids,  
 CC proteins, antipodes and (ant)agonists are useful in the diagnosis,  
 CC treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer,  
 CC and other cancers of the adrenal gland, bone, bone marrow, breast,  
 CC gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders  
 CC e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,  
 CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple  
 CC sclerosis, rheumatoid arthritis and ulcerative colitis;  
 CC (c) cardiovascular disorders such as myocardial ischaemia; (d) wound  
 CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;  
 CC and (f) infectious diseases such as viral, bacterial, fungal and  
 CC parasitic infections.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX  
 SQ Sequence 32192 BP; 8946 A; 6809 C; 6885 G; 9552 T; 0 other;  
 Query Match 90.0%; Score 18; DB 22; Length 32192;  
 Best Local Similarity 100.0%; Pred. No. 78;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 TTAGTCTTGTAGTCTTAG 18  
 |||||  
 Db 24031 TTAGTCTTGTAGTCTTAG 24014

Search completed: April 6, 2003, 12:05:51  
 Job time : 176 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 11:12:37 ; Search time 45.25 Seconds  
(without alignments)  
135.548 Million cell updates/sec

Title: US-09-980-265-11

Perfect score: 20

Sequence: 1 ttagctttagcttagctt 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA:\*

1: /cgn2\_6/ptodata/2/lna/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/lna/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/lna/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/lna/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/lna/PCTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/lna/Dackfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16.4	82.0	111282	4	US-09-754-250-3
2	15.8	79.0	1920	4	US-08-682-767-22
3	15.4	77.0	3162	2	US-08-449-645A-12
4	15.4	77.0	3162	2	US-08-702-367A-12
5	15.4	77.0	3162	5	PCN-US95-04681-12
6	15.2	76.0	995	4	US-08-976-259-101
7	15.2	76.0	1602	2	US-08-770-544-3
8	15.2	76.0	3231	1	US-08-195-152-1
9	15.2	76.0	5000	4	US-09-104-070-3
10	15.2	76.0	20165	4	US-09-609-816-7
11	15.2	76.0	59065	4	US-09-813-817-3
12	15.2	76.0	59065	4	US-09-978-197-3
13	15	75.0	1770	3	US-09-006-636-1
14	15	75.0	1770	4	US-09-006-636-1
15	15	75.0	1770	4	US-09-325-274-1
16	15	75.0	6000	3	US-09-006-636-9
17	15	75.0	6000	4	US-09-006-636-9
18	15	75.0	6000	4	US-09-325-274-9
19	14.8	74.0	524	4	US-09-222-575-169
20	14.8	74.0	96844	4	US-09-791-211-10
21	14.4	72.0	815	3	US-08-433-263B-1
22	14.4	72.0	2011	4	US-09-182-450-2
23	14.4	72.0	2181	4	US-09-780-049-17
24	14.4	72.0	2966	4	US-09-780-049-3
25	14.4	72.0	3381	3	US-08-937-195-1
26	14.4	72.0	3381	3	US-08-937-195-2
27	14.4	72.0	3381	4	US-08-915-152-1

C 28	14.4	72.0	3381	4	US-08-915-152-2	Sequence 2, Appl
C 29	14.4	72.0	3381	5	PCN-US96-07627-1	Sequence 1, Appl
C 30	14.4	72.0	3381	5	PCN-US96-07627-2	Sequence 2, Appl
C 31	14.4	72.0	40000	4	US-09-780-049-18	Sequence 18, Appl
C 32	14.2	71.0	228	1	US-08-222-177A-43	Sequence 43, Appl
C 33	14.2	71.0	362	4	US-09-605-785-442	Sequence 442, App
C 34	14.2	71.0	362	4	US-09-439-313-442	Sequence 442, App
C 35	14.2	71.0	362	4	US-09-352-616A-442	Sequence 46, Appl
C 36	14.2	71.0	427	4	US-09-556-877-46	Sequence 46, Appl
C 37	14.2	71.0	427	4	US-09-620-412C-46	Sequence 46, Appl
C 38	14.2	71.0	586	4	US-09-385-982-472	Sequence 472, App
C 39	14.2	71.0	1001	4	US-09-641-638-92	Sequence 92, Appl
C 40	14.2	71.0	1001	4	US-09-641-638-93	Sequence 93, Appl
C 41	14.2	71.0	1001	4	US-09-641-638-94	Sequence 94, Appl
C 42	14.2	71.0	1872	4	US-09-280-116-110	Sequence 110, App
C 43	14.2	71.0	1922	1	US-08-700-749A-7	Sequence 7, Appl
C 44	14.2	71.0	1922	3	US-09-020-684-7	Sequence 7, Appl
C 45	14.2	71.0	1922	3	US-09-020-467-7	Sequence 7, Appl

## ALIGNMENTS

```
RESULT 1
US-09-754-250-3
; Sequence 3, Application US/09754250
; Patent No. 6376225
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHODIESTERASE
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: C1001063
; CURRENT APPLICATION NUMBER: US/09/754,250
; CURRENT FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 111282
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)...(111282)
; OTHER INFORMATION: n = A,T,C or G
US-09-754-250-3

Query Match      82.0%; Score 16.4; DB 4; Length 111282;
Best Local Similarity 94.4%; Pred. No. 49;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      3 AGTCTTAGCTTAGT 20
        |||
Db 102136 AGTCTTAGCTTAGAT 102153

RESULT 2
US-08-682-767-22
; Sequence 22, Application US/08682767
; Patent No. 6291741
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: METHODS FOR THE PRODUCTION OF MODIFIED PLANTS
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of The Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/682,767  
 FILING DATE: 30-July-96  
 CLASSIFICATION: 800  
 ATTORNEY/AGENT INFORMATION:  
 NAME: White, John P  
 REGISTRATION NUMBER: 28,678  
 REFERENCE/DOCKET NUMBER: 51735  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-278-0400  
 TELEFAX: 212-391-0526  
 INFORMATION FOR SEQ ID NO: 22:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1920 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: unknown  
 MOLECULE TYPE: DNA (genomic)  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 US-08-682-767-22

Query Match 79.0%; Score 15.8; DB 4; Length 1920;  
 Best Local Similarity 89.5%; Pred. No. 76;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TAGTCTTAGTCTTAGT 20  
 Db 48 TATTCTTATTCTTAGT 66

RESULT 3  
 US-08-449-645A-12  
 Sequence 12, Application US/08449645A  
 Patent No. 5981245  
 GENERAL INFORMATION:  
 APPLICANT: Fox, Gary M.  
 TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine  
 TITLE OF INVENTION: Kinases  
 NUMBER OF SEQUENCES: 43  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Amgen Patent Operations/RBW  
 STREET: 1840 Dehaviiland Drive  
 CITY: Thousand Oaks  
 STATE: California  
 COUNTRY: USA  
 ZIP: 91320  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/449,645A  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Winter, Robert B.  
 REGISTRATION/DOCKET NUMBER: A-287  
 INFORMATION FOR SEQ ID NO: 12:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3162 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: CDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..2976

US-08-449-645A-12  
 Query Match 77.0%; Score 15.4; DB 2; Length 3162;  
 Best Local Similarity 94.1%; Pred. No. 1.1e+02;  
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 AGTCTTAGTCTTAGT 19  
 Db 3144 AGTCTTAGTCTTAGT 3160

RESULT 4  
 US-08-702-367A-12  
 Sequence 12, Application US/08702367A  
 Patent No. 5981246  
 GENERAL INFORMATION:  
 APPLICANT: Fox, Gary M.  
 TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine  
 TITLE OF INVENTION: Kinases  
 NUMBER OF SEQUENCES: 43  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Amgen Patent Operations/RBW  
 STREET: 1840 Dehaviiland Drive  
 CITY: Thousand Oaks  
 STATE: California  
 COUNTRY: USA  
 ZIP: 91320  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/702,367A  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Winter, Robert B.  
 REGISTRATION/DOCKET NUMBER: A-287  
 INFORMATION FOR SEQ ID NO: 12:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3162 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: CDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..2976  
 US-08-702-367A-12

Query Match 77.0%; Score 15.4; DB 2; Length 3162;  
 Best Local Similarity 94.1%; Pred. No. 1.1e+02;  
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 AGTCTTAGTCTTAGT 19  
 Db 3144 AGTCTTAGTCTTAGT 3160

RESULT 5  
 PCT-US95-04681-12  
 Sequence 12, Application PC/TUS9504681  
 GENERAL INFORMATION:  
 APPLICANT: Fox, Gary M.  
 TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine  
 TITLE OF INVENTION: Kinases  
 NUMBER OF SEQUENCES: 28  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Amgen Patent Operations/RBW  
 STREET: 1840 Dehaviiland Drive  
 CITY: Thousand Oaks  
 STATE: California



COUNTRY: USA  
ZIP: 91320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04681  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Winter, Robert B.  
REFERENCE/DOCKET NUMBER: A-287  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3162 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..2976  
PCT-US95-04681-12

Query Match 77.0% Score 15.4; DB 5; Length 3162;  
Best Local Similarity 94.1%; Pred. No. 1.1e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 ACTCTTACTTCTTAGT 19  
Db 3144 AGTCTTACTTCTTAGT 3160

RESULT 6  
US-08-976-259-101  
Sequence 101, Application US/08976259  
Patent No. 6316609  
GENERAL INFORMATION:  
APPLICANT: Dillon, Patrick J.  
APPLICANT: Choi, Gail H.  
APPLICANT: Welch, Rodney A.  
TITLE OF INVENTION: Nucleotide Sequence of Escherichia coli  
Patent No. 6316609  
NUMBER OF SEQUENCES: 142  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
STREET: 1100 New York Ave, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/976, 259  
FILING DATE: Herewith  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/031,626 AND US 60/061,953  
ATTORNEY/AGENT INFORMATION:  
NAME: Steffe, Eric K.  
REGISTRATION NUMBER: 36,688  
REFERENCE/DOCKET NUMBER: 1488.0740002/EKS/CM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 101:  
SEQUENCE CHARACTERISTICS:

LENGTH: 995 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-976-259-101

Query Match 76.0% Score 15.2; DB 4; Length 995;  
Best Local Similarity 85.0%; Pred. No. 1.3e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTAGTT 20  
Db 489 TTAGCTTACTGCTCTTAGCT 508

RESULT 7  
US-08-770-544-3  
Sequence 3, Application US/08770544  
Patent No. 5907085  
GENERAL INFORMATION:  
APPLICANT: Gonsalves, Dennis  
APPLICANT: Ling, Kai-Shu  
TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS PROTEINS AND  
THEIR USES  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
STREET: Clinton Square, P.O. Box 1051  
CITY: Rochester  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 14603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/770,544  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60009008  
FILING DATE: 21-DEC-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Goldman, Michael L.  
REGISTRATION NUMBER: 30,727  
REFERENCE/DOCKET NUMBER: 19603/621  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (716) 263-1304  
TELEFAX: (716) 263-1600  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1602 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-770-544-3

Query Match 76.0% Score 15.2; DB 2; Length 1602;  
Best Local Similarity 85.0%; Pred. No. 1.4e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTAGTT 20  
Db 1190 TTGTCTTAGTCTTAGTT 1209

RESULT 8  
US-08-195-152-1  
Sequence 1, Application US/08195152  
Patent No. 5679541



RESULT 12  
US-09-978-197-3  
; Sequence 3, Application US/09978197  
; Patent No. 6403353  
; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; FILE REFERENCE: C1001178DIV  
; CURRENT APPLICATION NUMBER: US/09/978,197  
; PRIORITY FILING DATE: 2001-10-17  
; PRIORITY APPLICATION NUMBER: 09/813,817  
; PRIORITY FILING DATE: 2001-03-22  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 59065  
; TYPE: DNA  
; ORGANISM: Human  
US-09-978-197-3

Query Match 76.0%; Score 15.2; DB 4; Length 59065;  
Best Local Similarity 85.0%; Pred. No. 1.5e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTTCTTACTT 20  
||| ||| ||| ||| ||| |||  
DB 15314 TTGTCTTCTTCTTACTT 15333

RESULT 13  
US-09-006-636-1/c  
; Sequence 1, Application US/09006636  
; Patent No. 6005092  
; GENERAL INFORMATION:  
; APPLICANT: Shoseyov, Oded  
; APPLICANT: Shani, Ziv  
; TITLE OF INVENTION: ARABIDOPSIS THALIANA ENDO-1,4-BETA-  
; TITLE OF INVENTION: GLUCANASE GENE, PROMOTER AND PROTEIN  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/006,636  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baldwin, Geraldine F.  
; REGISTRATION NUMBER: 31,232  
; REFERENCE/DOCKET NUMBER: 7809-019  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1770 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
US-09-006-636-1

Query Match 75.0%; Score 15; DB 3; Length 1770;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 AGTCTTACTTCTTA 17  
||||| ||| ||| ||| ||| |||  
DB 659 AGTCTTACTTCTTA 645

RESULT 14  
US-09-006-632-1/c  
; Sequence 1, Application US/09006632  
; Patent No. 6184440  
; GENERAL INFORMATION:  
; APPLICANT: Shoseyov, Oded  
; APPLICANT: Shani, Ziv  
; TITLE OF INVENTION: TRANSGENIC PLANTS OF ALTERED  
; TITLE OF INVENTION: MORPHOLOGY  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/006,632  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baldwin, Geraldine F.  
; REGISTRATION NUMBER: 31,232  
; REFERENCE/DOCKET NUMBER: 7809-018  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1770 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
US-09-006-632-1

Query Match 75.0%; Score 15; DB 4; Length 1770;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 AGTCTTACTTCTTA 17  
||||| ||| ||| ||| ||| |||  
DB 659 AGTCTTACTTCTTA 645

RESULT 15  
US-09-325-274-1/c  
; Sequence 1, Application US/09325274  
; Patent No. 6333023  
; GENERAL INFORMATION:  
; APPLICANT: Shoseyov, Oded  
; APPLICANT: Shani, Ziv  
; TITLE OF INVENTION: ARABIDOPSIS THALIANA ENDO-1,4-BETA-  
; TITLE OF INVENTION: GLUCANASE GENE, PROMOTER AND PROTEIN  
; NUMBER OF SEQUENCES: 21

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSO Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/325,274
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/006,636
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Baldwin, Geraldine F.
REGISTRATION NUMBER: 31,232
REFERENCE/DOCKET NUMBER: 7809-019
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1770 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-09-325-274-1

```

```

Query Match          75.0%; Score 15; DB 4; Length 1770;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 AGTCTTAGTCTTA 17
    |||||
Db 659 AGTCTTAGTCTTA 645

```

Search completed: April 6, 2003, 11:56:27  
 Job time : 87.25 secs

GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 11:24:32 ; Search time 58 Seconds  
(without alignments)  
302.472 Million cell updates/sec

Title: US-09-980-265-11

Perfect score: 20

Sequence: 1 ttacttcttagtcttagt 20

Scoring table: IDENTITY NUC

Searched: 593429 seqs, 438583890 residues

Total number of hits satisfying chosen parameters: 1186858

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*

2: /cgn2\_6/ptodata/1/pubpna/PC1\_NEM\_PUB.seq:\*

3: /cgn2\_6/ptodata/1/pubpna/US06\_NEM\_PUB.seq:\*

4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*

5: /cgn2\_6/ptodata/1/pubpna/US07\_NEM\_PUB.seq:\*

6: /cgn2\_6/ptodata/1/pubpna/PC105\_PUBCOMB.seq:\*

7: /cgn2\_6/ptodata/1/pubpna/US08\_NEM\_PUB.seq:\*

8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*

9: /cgn2\_6/ptodata/1/pubpna/US09\_NEM\_PUB.seq:\*

10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*

11: /cgn2\_6/ptodata/1/pubpna/US10\_NEM\_PUB.seq:\*

12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*

13: /cgn2\_6/ptodata/1/pubpna/US60\_NEM\_PUB.seq:\*

14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17	85.0	659158	9	US-09-771-208-20
2	16.8	84.0	548	10	US-09-924-035A-396
3	16.8	84.0	572	10	US-09-864-761-12035
4	16.8	84.0	4920	7	US-08-781-986A-26
5	16.4	82.0	111282	12	US-10-094-989-3
6	15.8	79.0	397	10	US-09-867-701-5873
7	15.4	77.0	338	10	US-09-878-574-5203
8	15.4	77.0	340	9	US-10-025-380-568
9	15.4	77.0	340	10	US-09-922-217-568
10	15.4	77.0	340	9	US-09-833-263-568
11	15.4	77.0	2000	9	US-09-938-842A-4252
12	15.2	76.0	207	10	US-09-867-701-4546
13	15.2	76.0	241	10	US-09-924-035A-72
14	15.2	76.0	471	9	US-09-736-457-524
15	15.2	76.0	471	9	US-09-902-941-524
16	15.2	76.0	471	9	US-09-849-626-524
17	15.2	76.0	471	9	US-10-017-754-524
18	15.2	76.0	483	9	US-10-202-193-239
19	15.2	76.0	711	10	US-09-770-149-200

20	15.2	76.0	742	10	US-09-770-149-51	Sequence 51, Appl
21	15.2	76.0	742	10	US-09-770-149-54	Sequence 54, Appl
22	15.2	76.0	815	9	US-10-202-193-240	Sequence 240, Appl
23	15.2	76.0	995	10	US-09-956-004-101	Sequence 101, Appl
24	15.2	76.0	1244	9	US-09-981-353-79	Sequence 79, Appl
25	15.2	76.0	1288	10	US-09-925-301-596	Sequence 596, Appl
26	15.2	76.0	1478	9	US-09-938-842A-2773	Sequence 2773, Appl
27	15.2	76.0	1557	9	US-10-001-873-1	Sequence 1, Appl
28	15.2	76.0	2000	9	US-09-938-842A-4454	Sequence 4454, Appl
29	15.2	76.0	2000	9	US-09-938-842A-4984	Sequence 4984, Appl
30	15.2	76.0	2000	9	US-09-938-842A-5278	Sequence 5278, Appl
31	15.2	76.0	2000	9	US-09-938-842A-5298	Sequence 5298, Appl
32	15.2	76.0	2002	10	US-09-887-576-3	Sequence 3, Appl
33	15.2	76.0	2010	10	US-09-887-576-2	Sequence 2, Appl
34	15.2	76.0	2785	10	US-09-802-674-3	Sequence 3, Appl
35	15.2	76.0	3231	8	US-08-754-311B-1	Sequence 1, Appl
36	15.2	76.0	3498	9	US-10-091-504-2204	Sequence 2204, Appl
37	15.2	76.0	3498	9	US-10-091-504-2271	Sequence 2271, Appl
38	15.2	76.0	3498	10	US-09-764-869-2204	Sequence 2204, Appl
39	15.2	76.0	3498	10	US-09-764-869-2271	Sequence 2271, Appl
40	15.2	76.0	3499	9	US-10-091-504-2203	Sequence 2203, Appl
41	15.2	76.0	3499	9	US-10-091-504-2270	Sequence 2270, Appl
42	15.2	76.0	3499	10	US-09-764-869-2203	Sequence 2203, Appl
43	15.2	76.0	3499	10	US-09-764-869-2270	Sequence 2270, Appl
44	15.2	76.0	20165	9	US-10-199-330-7	Sequence 7, Appl
45	15.2	76.0	20165	9	US-10-199-334-7	Sequence 7, Appl

## ALIGNMENTS

RESULT 1

US-09-771-208-20

Sequence 20, Application US/09771208

Patent No. US20020155564A1

GENERAL INFORMATION:

APPLICANT: MEDRANO, JUAN

APPLICANT: BRADFORD, ERIC

APPLICANT: HORVAT, SIMON

TITLE OF INVENTION: CLONING OF A HIGH-GROWTH GENE

FILE REFERENCE: 407T-923710US

CURRENT FILING DATE: 2001-01-26

PRIOR APPLICATION NUMBER: US 08/999,477

PRIOR FILING DATE: 1997-12-29

NUMBER OF SEQ ID NOS: 20

SOFTWARE: PatentIn version 3.0

SEQ ID NO 20

LENGTH: 659158

TYPE: DNA

ORGANISM: Mus musculus

FEATURE:

NAME/KEY: misc.feature

LOCATION: (123459)..(123478)

OTHER INFORMATION: n is unidentified a, c, g, or t

NAME/KEY: misc.feature

LOCATION: (602466)..(602485)

OTHER INFORMATION: n is unidentified a, c, g, or t

NAME/KEY: misc.feature

LOCATION: (546598)..(547017)

OTHER INFORMATION: n is unidentified a, c, g, or t

NAME/KEY: misc.feature

LOCATION: (494715)..(494814)

OTHER INFORMATION: n is unidentified a, c, g, or t

NAME/KEY: misc.feature

LOCATION: (390866)..(391005)

OTHER INFORMATION: n is unidentified a, c, g, or t

NAME/KEY: misc.feature

LOCATION: (346860)..(346823)

OTHER INFORMATION: n is unidentified a, c, g, or t

NAME/KEY: misc.feature

LOCATION: (317174)..(317193)

OTHER INFORMATION: n is unidentified a, c, g, or t

```
NAME/KEY: misc_feature
LOCATION: (280353)..(280373)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (271829)..(271848)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (183872)..(183891)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (170625)..(170645)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (132680)..(132700)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
OTHER INFORMATION: n is a, c, g, or t
US-09-771-208-20
```

```
Query Match      85.0%; Score 17; DB 9; Length 659158;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      4 GTTCTAGTCTTAGT 20
Db 375807 GTTCTAGTCTTAGT 375823
```

```
RESULT 2
US-09-924-035A-396/c
Sequence 396, Application US/09924035A
Patent No. US20020142319A1
GENERAL INFORMATION:
APPLICANT: GliaCh, Jrn
TITLE OF INVENTION: Expressed Sequences of Arabidopsis
FILE REFERENCE: 2011US
CURRENT APPLICATION NUMBER: US/09/924,035A
CURRENT FILING DATE: 2000-08-11
PRIOR APPLICATION NUMBER: US 60/148,784
PRIOR FILING DATE: 1999-08-13
NUMBER OF SEQ ID NOS: 900
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 396
LENGTH: 548
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(348)
OTHER INFORMATION: n = A,T,C or G
US-09-924-035A-396
```

```
Query Match      84.0%; Score 16.8; DB 10; Length 548;
Best Local Similarity 90.0%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY      1 TTAGTCTTAGTCTTAGT 20
Db 172 TTAGTCTTAGTCTTAGT 153
```

```
RESULT 3
US-09-864-761-12035
Sequence 12035, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
GENE EXPRESSION ANALYSIS BY MICROARRAY
```

```
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 12035
LENGTH: 572
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL133372.2
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.83
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.66
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.7
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.64
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.6
US-09-864-761-12035
```

```
Query Match      84.0%; Score 16.8; DB 10; Length 572;
Best Local Similarity 90.0%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY      1 TTAGTCTTAGTCTTAGT 20
Db 403 TTATTCTTAGTCTTAGT 422
```

```
RESULT 4
US-08-781-986A-26/c
Sequence 26, Application US/08781986A
Publication No. US20030054436A1
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
```

STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/781.986A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Benson, Bob  
REGISTRATION NUMBER: 30,446  
REFERENCE/DOCKET NUMBER: PB248BP  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4920 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-781-986A-26

Query Match 84.0%; Score 16.8; DB 7; Length 4920;  
Best Local Similarity 90.0%; Pred. No. 1.6e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTAGTT 20  
||| ||||| ||||| |||||  
DB 564 TTAATCTTACTTCTTAGTT 545

RESULT 5  
US-10-094-989-3  
Sequence 3, Application US/10094989  
Patent No. US20020115179A1  
GENERAL INFORMATION:  
APPLICANT: WEI, Ming-Hui et al  
TITLE OF INVENTION: ISOLATED HUMAN PHOSPHODIESTERASE  
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN  
FILE REFERENCE: CL001063DIY  
CURRENT APPLICATION NUMBER: US/10/094,989  
CURRENT FILING DATE: 2002-03-12  
PRIOR APPLICATION NUMBER: 09/754,250  
PRIOR FILING DATE: 2001-01-05  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 111282  
TYPE: DNA  
ORGANISM: Homo sapien  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...((111282)  
OTHER INFORMATION: n = A,T,C or G  
US-10-094-989-3

Query Match 82.0%; Score 16.4; DB 12; Length 111282;  
Best Local Similarity 94.4%; Pred. No. 3.3e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 AGTTCTTAGTCTTAGTT 20  
||||| ||||| ||||| |||||

DB 102136 AGTCTTAGTCTTAGAT 102153

RESULT 6  
US-09-867-701-5873  
Sequence 5873, Application US/09867701  
Patent No. US20020132237A1  
GENERAL INFORMATION:  
APPLICANT: Aglate, Paul A.  
APPLICANT: Jones, Robert  
APPLICANT: Harlocker, Susan L.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
FILE REFERENCE: 210121.497  
CURRENT APPLICATION NUMBER: US/09/867,701  
CURRENT FILING DATE: 2001-05-29  
NUMBER OF SEQ ID NOS: 10912  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5873  
LENGTH: 397  
TYPE: DNA  
ORGANISM: Homo sapien  
US-09-867-701-5873

Query Match 79.0%; Score 15.8; DB 10; Length 397;  
Best Local Similarity 89.5%; Pred. No. 3.2e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TAGTCTTAGTCTTAGTT 20  
||| ||||| ||||| |||||  
DB 100 TAGTCTTAGTCTTAGAT 118

RESULT 7  
US-09-878-574-5203  
Sequence 5203, Application US/09878574  
Patent No. US20020110548A1  
GENERAL INFORMATION:  
APPLICANT: Byrum, Joseph R.  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Thompson, Michael D.  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants  
FILE REFERENCE: 38-21(15401)B  
CURRENT APPLICATION NUMBER: US/09/878,574  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 09/333,535  
PRIOR FILING DATE: 1999-06-14  
NUMBER OF SEQ ID NOS: 15775  
SEQ ID NO 5203  
LENGTH: 338  
TYPE: DNA  
ORGANISM: Glycine max  
OTHER INFORMATION: Clone ID: LIB3028-029-01-B1-E5  
US-09-878-574-5203

Query Match 77.0%; Score 15.4; DB 10; Length 338;  
Best Local Similarity 94.1%; Pred. No. 4.6e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTAGTT 17  
||| ||||| ||||| |||||  
DB 215 TTAATCTTACTTCTTAGTT 231

RESULT 8  
US-10-025-380-568  
Sequence 568, Application US/10025380  
Publication No. US20020182191A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Lodes, Michael J.  
APPLICANT: Secrist, Heather

APPLICANT: Benson, Darin R.  
APPLICANT: Meagher, Madeleine Joy  
APPLICANT: Stolk, John A.  
APPLICANT: Wang, Tongtong  
APPLICANT: Jiang, Yugu  
APPLICANT: Smith, Carole L.  
APPLICANT: King, Gordon E.  
APPLICANT: Wang, Aljun  
APPLICANT: Clapper, Jonathan D.  
APPLICANT: Skelky, Yasir A. W.  
APPLICANT: Fanger, Gary R.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darick  
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS  
FILE REFERENCE: 210121.471C14  
CURRENT APPLICATION NUMBER: US/10/025.380  
CURRENT FILING DATE: 2001-12-19  
NUMBER OF SEQ ID NOS: 1129  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 568  
LENGTH: 340  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 326  
OTHER INFORMATION: n = A,T,C or G  
US-10-025-380-568

Query Match 77.0% Score 15.4; DB 9; Length 340;  
Best Local Similarity 94.1% Pred. No. 4.6e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTAGTCTTAGTCTTA 17  
Db 227 TTGGTCTTAGTCTTA 243

RESULT 9  
US-09-922-217-568  
Sequence 568, Application US/09922217  
Patent No. US20020076414A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Lodes, Michael J.  
APPLICANT: Secrist, Heather  
APPLICANT: Benson, Darin R.  
APPLICANT: Meagher, Madeleine Joy  
APPLICANT: Stolk, John A.  
APPLICANT: Wang, Tongtong  
APPLICANT: Jiang, Yugu  
APPLICANT: Smith, Carole Lynn  
APPLICANT: King, Gordon E.  
APPLICANT: Wang, Aljun  
APPLICANT: Clapper, Jonathan D.  
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS  
FILE REFERENCE: 210121.471C13  
CURRENT APPLICATION NUMBER: US/09/922.217  
CURRENT FILING DATE: 2001-08-03  
NUMBER OF SEQ ID NOS: 1124  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 568  
LENGTH: 340  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 326  
OTHER INFORMATION: n = A,T,C or G  
US-09-922-217-568

Query Match 77.0% Score 15.4; DB 10; Length 340;  
Best Local Similarity 94.1% Pred. No. 4.6e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 TTAGTCTTAGTCTTA 17  
Db 227 TTGGTCTTAGTCTTA 243

RESULT 10  
US-09-833-263-568  
Sequence 568, Application US/09833263  
Patent No. US2002010547A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Aljun  
APPLICANT: Clapper, Jonathan D.  
APPLICANT: Stolk, John A.  
APPLICANT: Meagher, Madeleine J.  
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND  
FILE REFERENCE: 210121.471C12  
CURRENT APPLICATION NUMBER: US/09/833.263  
CURRENT FILING DATE: 2001-04-10  
NUMBER OF SEQ ID NOS: 1093  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 568  
LENGTH: 340  
TYPE: DNA  
ORGANISM: Homo sapien  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(340)  
OTHER INFORMATION: n = A,T,C or G  
US-09-833-263-568

Query Match 77.0% Score 15.4; DB 10; Length 340;  
Best Local Similarity 94.1% Pred. No. 4.6e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTAGTCTTAGTCTTA 17  
Db 227 TTGGTCTTAGTCTTA 243

RESULT 11  
US-09-938-842A-4252  
Sequence 4252, Application US/09938842A  
Patent No. US20020160378A1  
GENERAL INFORMATION:  
APPLICANT: Harper, Jeff  
APPLICANT: Kreps, Joel  
APPLICANT: Wang, Xun  
APPLICANT: Zhu, Tong  
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
FILE REFERENCE: SRIPI300-3  
CURRENT APPLICATION NUMBER: US/09/938.842A  
CURRENT FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: US 60/227,866  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: US 60/264,647  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/300,111  
PRIOR FILING DATE: 2001-06-22  
NUMBER OF SEQ ID NOS: 5379  
SEQ ID NO 4252  
LENGTH: 2000  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-938-842A-4252

Query Match 77.0% Score 15.4; DB 9; Length 2000;  
Best Local Similarity 94.1% Pred. No. 5.7e+02;



Matches 16: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

QY 1 TTAGTCTTAGTCTTCTTA 17  
|||||  
Db 1367 TTAGTCTTAGTCTTCTTA 1383

RESULT 12  
US-09-867-701-4546/C  
; Sequence 4546, Application US/09867701  
; Patent No. US20020132237A1  
; GENERAL INFORMATION:  
; APPLICANT: Aglate, Paul A.  
; APPLICANT: Jones, Robert  
; APPLICANT: Harlocker, Susan L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.497  
; CURRENT APPLICATION NUMBER: US/09/867,701  
; CURRENT FILING DATE: 2001-05-29  
; NUMBER OF SEQ ID NOS: 10912  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4546  
; LENGTH: 207  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-867-701-4546

Query Match 76.0%; Score 15.2; DB 10; Length 207;  
Best Local Similarity 85.0%; Pred. No. 5.3e+02;  
Matches 17: Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTCTTCTT 20  
|||||  
Db 113 TTGTTCTTAGTCTTCTTCTT 94

RESULT 13  
US-09-924-035A-72/C  
; Sequence 72, Application US/09924035A  
; Patent No. US20020142319A1  
; GENERAL INFORMATION:  
; APPLICANT: Grilach, Jrn  
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
; FILE REFERENCE: 2011US  
; CURRENT APPLICATION NUMBER: US/09/924,035A  
; CURRENT FILING DATE: 2000-08-11  
; PRIOR APPLICATION NUMBER: US 60/148,784  
; PRIOR FILING DATE: 1999-08-13  
; NUMBER OF SEQ ID NOS: 900  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 72  
; LENGTH: 241  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-924-035A-72

Query Match 76.0%; Score 15.2; DB 10; Length 241;  
Best Local Similarity 85.0%; Pred. No. 5.4e+02;  
Matches 17: Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTCTTCTT 20  
|||||  
Db 87 TTACTTCTTAATCTTATTT 68

RESULT 14  
US-09-736-457-524/C  
; Sequence 524, Application US/09736457  
; Patent No. US20020168637A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tonglong

; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darlick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; APPLICANT: Wang, Aijun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.478C15  
; CURRENT APPLICATION NUMBER: US/09/736,457  
; CURRENT FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 1864  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 524  
; LENGTH: 471  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-736-457-524

Query Match 76.0%; Score 15.2; DB 9; Length 471;  
Best Local Similarity 85.0%; Pred. No. 5.9e+02;  
Matches 17: Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTCTTCTT 20  
|||||  
Db 190 TTAGCTTTAGTCTTGTGTT 171

RESULT 15  
US-09-902-941-524/C  
; Sequence 524, Application US/09902941  
; Patent No. US20020172952A1  
; GENERAL INFORMATION:  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Wang, Tonglong  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Johnson, Jeffrey C.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Marnerakis, Margarita  
; APPLICANT: Carter, Darlick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: McNabb, Andria  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.478C17  
; CURRENT APPLICATION NUMBER: US/09/902,941  
; CURRENT FILING DATE: 2001-07-10  
; NUMBER OF SEQ ID NOS: 2002  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 524  
; LENGTH: 471  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-902-941-524

Query Match 76.0%; Score 15.2; DB 9; Length 471;  
Best Local Similarity 85.0%; Pred. No. 5.9e+02;  
Matches 17: Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTCTTCTT 20  
|||||  
Db 190 TTAGCTTTAGTCTTGTGTT 171

Search completed: April 6, 2003, 12:12:31  
Job time : 212 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 10:29:22 ; Search time 168.5 seconds  
(without alignments)  
1912.600 Million cell updates/sec

Title: US-09-980-265-12  
Perfect score: 20  
Sequence: 1 ttagttgttagttattgctt 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 60155 seqs, 8056825618 residues

Total number of hits satisfying chosen parameters: 120310

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl\_HTG:\*  
1: gb\_hcgc:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	19.4	95.0	195444	AC107399	AC107399 Homo sapi
2	18.4	92.0	5421	AC019958	AC019958 Drosophila
3	18.4	92.0	38069	AC017255	AC017255 Drosophila
4	18.4	92.0	44765	AC116467	AC116467 Mus muscu
5	18.4	92.0	80679	AC103576	AC103576 Rattus no
6	18.4	92.0	95180	AC109057	AC109057 Rattus no
7	18.4	92.0	106651	AC106466	AC106466 Rattus no
8	18.4	92.0	150326	AC128956	AC128956 Rattus no
9	18.4	92.0	152324	AC129017	AC129017 Mus muscu
10	18.4	92.0	164552	AC111380	AC111380 Rattus no
11	18.4	92.0	176400	AC113369	AC113369 Homo sapi
12	18.4	92.0	178712	AC115176	AC115176 Rattus no
13	18.4	92.0	178970	AC012589	AC012589 Homo sapi
14	18.4	92.0	201132	AC122335	AC122335 Mus muscu
15	18.4	92.0	241585	AL844536	AL844536 Mus muscu
16	17.4	87.0	20733	AC019829	AC019829 Drosophila
17	17.4	87.0	60625	AC020090	AC020090 Drosophila
18	17.4	87.0	64061	AC100401	AC100401 Mus muscu
19	17.4	87.0	92053	AC017270	AC017270 Drosophila
20	17.4	87.0	154354	AC017270	AC017270 Drosophila
21	17.4	87.0	167683	AC079251	AC079251 Homo sapi
22	17.4	87.0	171272	AC084707	AC084707 Homo sapi
23	17.4	87.0	171457	AC112568	AC112568 Rattus no
24	17.4	87.0	174877	AP004818	AP004818 Oryza sat
25	17.4	87.0	177741	AC118176	AC118176 Rattus no
26	17.4	87.0	180485	AC099106	AC099106 Rattus no
27	17.4	87.0	183450	AC119322	AC119322 Rattus no
28	17.4	87.0	186385	AC026821	AC026821 Homo sapi
29	17.4	87.0	186542	AL844197	AL844197 Homo sapi
30	17.4	87.0	193039	AC090690	AC090690 Homo sapi
31	17.4	87.0	195600	AC079094	AC079094 Homo sapi
32	17.4	87.0	209572	AC099615	AC099615 Mus muscu

C 33	17.4	87.0	209790	1	AL512373	Homo sapi
C 34	17.4	87.0	216464	1	AC102540	Mus muscu
C 35	17.4	87.0	247102	1	AC123854	Mus muscu
C 36	17.4	87.0	267052	1	AC122872	Mus muscu
C 37	17.4	85.0	22154	1	AC115602	Dicystose
C 38	17.4	85.0	56759	1	AC025379	Homo sapi
C 39	17.4	85.0	68559	1	AC031981	Homo sapi
C 40	17.4	85.0	160294	1	AC025621	Homo sapi
C 41	17.4	85.0	166349	1	AC013707	Homo sapi
C 42	17.4	85.0	168208	1	AC024359	Homo sapi
C 43	17.4	85.0	214271	1	AC090737	Homo sapi
C 44	16.8	84.0	8659	1	AC117270	Dicystose
C 45	16.8	84.0	14921	1	AC015356	Drosophila

## ALIGNMENTS

RESULT 1  
AC107399/c  
LOCUS  
DEFINITION Homo sapiens chromosome 4 clone RP11-755N14, WORKING DRAFT  
ACCESSION AC107399.5 GI:21450604  
VERSION AC107399.5  
KEYWORDS HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ACTIVEPIN.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
Waterston, R.H.  
1 (bases 1 to 195444)  
2 (bases 1 to 195444)  
The sequence of Homo sapiens clone  
Unpublished  
Waterston, R.H.  
Direct Submission  
Submitted (19-JUN-2002) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
3 (bases 1 to 195444)  
Waterston, R.H.  
Direct Submission  
Submitted (20-JUN-2002) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
On Jun 18, 2002 this sequence version replaced gi:21426762.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc/index.shtml  
Contact: submissions@wustl.edu  
Project Information  
Center project name: H.NH0755N14

----- Summary Statistics -----  
Sequencing vector: M13, 0%  
Sequencing vector: Plasmid, 100%  
Chemistry: Dye-terminator Big Dye, 100% of reads  
Chemistry: Dye-terminator Big Dye, 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 194367 bases at least Q40  
Consensus quality: 195506 bases at least Q30  
Consensus quality: 195506 bases at least Q20  
Insert size: 211000; agarose-fp  
Insert size: 196567; sum-of-ctrls  
Quality coverage: 7.60 in Q20 bases; sum-of-ctrls  
Quality coverage: 6.70 in Q20 bases; sum-of-ctrls

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 3 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is



TITLE  
JOURNAL  
COMMENT

Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,  
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
Strauss, N., Sudramanian, A., Talamas, J., Testaye, S., Theodore, J.,  
Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Submitted (28-MAR-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/rw/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L23458

Center clone name: 165\_I\_15

\* NOTE: This record contains 58 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1  
\* 655 754: gap of 100 bp in length  
\* 755 1457: contig of 703 bp in length  
\* 1458 1557: gap of 100 bp  
\* 1558 2210: contig of 653 bp in length  
\* 2211 2310: gap of 100 bp  
\* 2311 2988: contig of 678 bp in length  
\* 2989 3088: gap of 100 bp  
\* 3089 3776: contig of 688 bp in length  
\* 3777 3876: gap of 100 bp  
\* 3877 4535: contig of 659 bp in length  
\* 4536 4635: gap of 100 bp  
\* 4636 5329: contig of 694 bp in length  
\* 5330 5429: gap of 100 bp  
\* 5430 6104: contig of 675 bp in length  
\* 6105 6204: gap of 100 bp  
\* 6205 6879: contig of 675 bp in length  
\* 6880 6979: gap of 100 bp  
\* 6980 7669: contig of 690 bp in length  
\* 7670 7769: gap of 100 bp  
\* 7770 8440: contig of 671 bp in length  
\* 8441 8540: gap of 100 bp  
\* 8541 9219: contig of 679 bp in length  
\* 9220 9319: gap of 100 bp  
\* 9320 9990: contig of 671 bp in length  
\* 9991 10090: gap of 100 bp  
\* 10091 10746: contig of 656 bp in length  
\* 10747 10846: gap of 100 bp  
\* 10847 11497: contig of 651 bp in length  
\* 11498 11597: gap of 100 bp  
\* 11598 12303: contig of 706 bp in length  
\* 12304 12403: gap of 100 bp  
\* 12404 13046: contig of 643 bp in length  
\* 13047 13146: gap of 100 bp  
\* 13147 13834: contig of 688 bp in length  
\* 13835 13934: gap of 100 bp  
\* 13935 14610: contig of 676 bp in length  
\* 14611 14710: gap of 100 bp  
\* 14711 15396: contig of 686 bp in length

15397 15496: gap of 100 bp  
15497 16152: contig of 656 bp in length  
16153 16252: gap of 100 bp  
16253 16949: contig of 697 bp in length  
16950 17049: gap of 100 bp  
17050 17720: contig of 671 bp in length  
17721 17820: gap of 100 bp  
17821 18524: contig of 704 bp in length  
18525 18624: gap of 100 bp  
18625 19335: contig of 711 bp in length  
19336 19435: gap of 100 bp  
19436 20126: contig of 691 bp in length  
20127 20226: gap of 100 bp  
20227 20906: contig of 680 bp in length  
20907 21006: gap of 100 bp  
21007 21687: contig of 681 bp in length  
21688 21787: gap of 100 bp  
21788 22456: contig of 669 bp in length  
22457 22556: gap of 100 bp  
22557 23229: contig of 673 bp in length  
23230 23329: gap of 100 bp  
23330 24011: contig of 682 bp in length  
24012 24111: gap of 100 bp  
24112 24758: contig of 647 bp in length  
24759 24858: gap of 100 bp  
24859 25531: contig of 673 bp in length  
25532 25631: gap of 100 bp  
25632 26304: contig of 673 bp in length  
26305 26404: gap of 100 bp  
26405 27034: contig of 630 bp in length  
27035 27134: gap of 100 bp  
27135 27841: contig of 707 bp in length  
27842 27941: gap of 100 bp  
27942 28584: contig of 643 bp in length  
28585 28684: gap of 100 bp  
28685 29346: contig of 662 bp in length  
29347 29446: gap of 100 bp  
29447 30136: contig of 690 bp in length  
30137 30236: gap of 100 bp  
30237 30906: contig of 670 bp in length  
30907 31006: gap of 100 bp  
31007 31686: contig of 680 bp in length  
31687 31786: gap of 100 bp  
31787 32456: contig of 670 bp in length  
32457 32556: gap of 100 bp  
32557 33196: contig of 640 bp in length  
33197 33286: gap of 100 bp  
33297 33977: contig of 681 bp in length  
33978 34077: gap of 100 bp  
34078 34704: contig of 627 bp in length  
34705 34804: gap of 100 bp  
34805 35481: contig of 677 bp in length  
35482 35581: gap of 100 bp  
35582 36251: contig of 670 bp in length  
36252 36351: gap of 100 bp  
36352 37049: contig of 698 bp in length  
37050 37149: gap of 100 bp  
37150 37997: contig of 648 bp in length  
37998 37887: gap of 100 bp  
37898 38601: contig of 704 bp in length  
38602 38701: gap of 100 bp  
38702 39324: contig of 623 bp in length  
39325 39424: gap of 100 bp  
39425 40112: contig of 688 bp in length  
40113 40212: gap of 100 bp  
40213 40892: contig of 680 bp in length  
40893 40992: gap of 100 bp  
40993 41656: contig of 664 bp in length  
41657 41756: gap of 100 bp  
41757 42447: contig of 691 bp in length  
42448 42547: gap of 100 bp  
42548 43206: contig of 659 bp in length  
43207 43306: gap of 100 bp

```

FEATURES
  source
    * 43307 43984: contig of 678 bp in length
    * 43985 44084: gap of 100 bp
    * 44085 44765: contig of 681 bp in length.
      location/qualifiers
      1..44765
      /organism="Mus musculus"
      /db_xref="taxon:10090"
      /clone="RP23-165115"
      /clone_id="RPC1-23 Female Mouse BAC"

BASE COUNT  11666 a 7677 c 7735 g 11806 t 5881 others
ORIGIN
  Query Match 92.0% Score 18.4; DB 1; Length 44765;
  Best Local Similarity 95.0%; Pred. No. 17;
  Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTACTGTCTTACTTACTT 20
   |||||  |||||  |||||
Db 4039 TTACTGTCTTACTTACTT 4058

RESULT 5
AC103576
LOCUS
DEFINITION Rattus norvegicus clone CH230-63M13, *** SEQUENCING IN PROGRESS
AC103576
VERSION AC103576.4 GI:21729709
KEYWORDS HTG; HTGS; PHASEL.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
          1 (bases 1 to 80679)
          Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
          Alshrooks,S.L., Amaralunga,H.C., Are,J.R., Ayele,M., Banks,T.,
          Barbadaia,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
          Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
          Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
          Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
          Chen,G., Chen,R., Chen,Z., Chowhry,I., Christopoulos,C.,
          Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
          Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
          Delaune,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
          Doulatwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
          Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
          Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
          Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
          Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
          Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
          Hernandez,O., Hodgson,A., Hogues,M., Hollway,C., Hollins,B.,
          Homsl,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
          Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
          Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
          Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
          Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,M., Louised,H.,
          Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
          Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
          Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
          Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
          Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
          Nguyen,N., Nickerson,E., Nwokwenkwo,S., Oguh,M., Okwunodu,G.,
          Otagunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
          Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
          Rivers,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Severy,G.,
          Scherer,S., Scott,G., Shen,H., Shooshari,N., Sisson,I.,
          Sodergren,E., Sonalike,T., Sparks,A., Stanley,H., Stone,H.,
          Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
          Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
          Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
          Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watling,C.,
          Williams,G., Williamson,A., Wleczyk,R., Wooden,C., Worley,K.,

```

```

REFERENCE
  TITLE
  JOURNAL
  REFERENCE
  AUTHORS
  TITLE
  JOURNAL
  REFERENCE
  AUTHORS
  TITLE
  JOURNAL
  COMMENT

  Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
  Weinstein,G., and Gibbs,R.
  Direct Submission
  Unpublished
  2 (bases 1 to 80679)
  Morley,K.C.
  Direct Submission
  Submitted (29-NOV-2001) Human Genome Sequencing Center, Department
  of Molecular and Human Genetics, Baylor College of Medicine, One
  Baylor Plaza, Houston, TX 77030, USA
  3 (bases 1 to 80679)
  Morley,K.C.
  Direct Submission
  Submitted (12-JUL-2002) Human Genome Sequencing Center, Department
  of Molecular and Human Genetics, Baylor College of Medicine, One
  Baylor Plaza, Houston, TX 77030, USA
  On Jul 11, 2002 this sequence version replaced gi:18846139.

  ----- Genome Center
  Center: Baylor College of Medicine
  Center code: BCM
  Web site: http://www.hgsc.bcm.tmc.edu/
  Contact: hgsc-help@bcm.tmc.edu
  ----- Project Information
  Center project name: GIGY
  Center clone name: CH230-63M13
  ----- Summary Statistics
  Chemistry: Dye-terminator Big Dye; 100% of reads
  Assembly program: Phrap; version 0.990329
  Consensus quality: 33297 bases at least 940
  Consensus quality: 35946 bases at least 930
  Consensus quality: 38652 bases at least 920
  -----
  * NOTE: Estimated insert size may differ from sequence length
  * (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
  * NOTE: This is a 'working draft' sequence. It currently
  * consists of 47 contigs. The true order of the pieces
  * is not known and their order in this sequence record is
  * arbitrary. Gaps between the contigs are represented as
  * runs of N, but the exact sizes of the gaps are unknown.
  * This record will be updated with the finished sequence
  * as soon as it is available and the accession number will
  * be preserved.
  1
  1171: contig of 1171 bp in length
  1172: gap of unknown length
  1272: contig of 1597 bp in length
  2868: gap of unknown length
  2969: contig of 1163 bp in length
  4131: gap of unknown length
  4132: gap of unknown length
  4232: contig of 1081 bp in length
  5313: gap of unknown length
  5413: contig of 1458 bp in length
  6870: gap of unknown length
  6971: contig of 1021 bp in length
  7991: gap of unknown length
  8091: gap of unknown length
  8092: contig of 1035 bp in length
  9127: gap of unknown length
  9227: gap of unknown length
  10468: contig of 1242 bp in length
  10569: gap of unknown length
  11619: contig of 1051 bp in length
  11719: gap of unknown length
  11720: gap of unknown length
  12894: contig of 1175 bp in length
  12895: gap of unknown length
  14178: contig of 1184 bp in length
  14278: gap of unknown length
  14279: contig of 1337 bp in length
  15615: gap of unknown length
  15716: gap of unknown length
  17017: contig of 1301 bp in length
  17116: gap of unknown length
  18233: contig of 1117 bp in length
  18234: gap of unknown length
  18334: contig of 1332 bp in length
  19665: gap of unknown length
  19766: contig of 1386 bp in length
  21151:

```

\* 21152 21251: gap of unknown length  
\* 21252 22739: contig of 1488 bp in length  
\* 22740 22839: gap of unknown length  
\* 22840 23909: contig of 1070 bp in length  
\* 23910 24009: gap of unknown length  
\* 24010 25502: contig of 1493 bp in length  
\* 25503 25602: gap of unknown length  
\* 25603 26796: contig of 1194 bp in length  
\* 26797 26896: gap of unknown length  
\* 26897 28120: contig of 1224 bp in length  
\* 28121 28220: gap of unknown length  
\* 28221 29944: contig of 1724 bp in length  
\* 29945 30044: gap of unknown length  
\* 30045 32238: contig of 2194 bp in length  
\* 32239 32338: gap of unknown length  
\* 32339 33569: contig of 1231 bp in length  
\* 33570 33669: gap of unknown length  
\* 33670 35560: contig of 1891 bp in length  
\* 35561 35660: gap of unknown length  
\* 35661 36788: contig of 1128 bp in length  
\* 36789 36888: gap of unknown length  
\* 36889 38239: contig of 1351 bp in length  
\* 38240 38339: gap of unknown length  
\* 38340 39574: contig of 1235 bp in length  
\* 39575 39674: gap of unknown length  
\* 39675 40719: contig of 1045 bp in length  
\* 40720 40819: gap of unknown length  
\* 40820 42444: contig of 1625 bp in length  
\* 42445 42544: gap of unknown length  
\* 42545 44516: contig of 1972 bp in length  
\* 44517 44616: gap of unknown length  
\* 44617 46392: contig of 1776 bp in length  
\* 46393 46492: gap of unknown length  
\* 46493 47848: contig of 1356 bp in length  
\* 47849 47948: gap of unknown length  
\* 47949 49380: contig of 1432 bp in length  
\* 49381 49480: gap of unknown length  
\* 49481 50992: contig of 1512 bp in length  
\* 50993 51092: gap of unknown length  
\* 51093 53214: contig of 2122 bp in length  
\* 53215 53314: gap of unknown length  
\* 53315 54686: contig of 1372 bp in length  
\* 54687 54786: gap of unknown length  
\* 54787 56476: contig of 1690 bp in length  
\* 56477 56576: gap of unknown length  
\* 56577 58624: contig of 2048 bp in length  
\* 58625 58724: gap of unknown length  
\* 58725 61311: contig of 2587 bp in length  
\* 61312 61411: gap of unknown length  
\* 61412 63407: contig of 1996 bp in length  
\* 63408 63507: gap of unknown length  
\* 63508 65412: contig of 1905 bp in length  
\* 65413 65512: gap of unknown length  
\* 65513 68121: contig of 2609 bp in length  
\* 68122 68221: gap of unknown length  
\* 68222 71194: contig of 2973 bp in length  
\* 71195 71294: gap of unknown length  
\* 71295 73654: contig of 2360 bp in length  
\* 73655 73754: gap of unknown length  
\* 73755 76459: contig of 2705 bp in length  
\* 76460 76559: gap of unknown length  
\* 76560 80679: contig of 4120 bp in length.

## FEATURES

Location/Qualifiers

1. 80679  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/clone="CH230-63N13"  
BASE COUNT 17902 a 18300 c 17407 g 18485 t 8585 others  
ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 80679;  
Best Local Similarity 95.0%; Pred. No. 18;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTGTAGTTATAGTT 20  
||||| |||||||||  
Db 41343 TTAGTTATAGTTATAGTT 41362

RESULT 6  
AC109057/c  
LOCUS  
DEFINITION  
Rattus norvegicus clone CH230-126F19, \*\*\* SEQUENCING IN PROGRESS  
\*\*\*48 unordered pieces.  
AC109057 95180 bp DNA linear HTG 13-JUL-2002  
AC109057  
AC109057.3 GI:21737841  
VERSION  
KEYWORDS  
HTG; HTGS\_PHASE1.  
SOURCE  
ORGANISM  
Norway rat.  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Ratus.

## REFERENCE

1 (bases 1 to 95180)  
Muzny,D.M., Adams,C., Adio-Oduola,B., Alt-osman,F.R., Allen,C.,  
Alshrocks,S.L., Amaratunga,H.C., Are,J.R., Ayala,M., Banks,T.,  
Barbaita,U., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,  
Bouck,U., Bowle,S., Brlewa,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Eamhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Correll,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
Homsli,F., Howard,S., Huber,J., Huijk,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,Y., Louisedge,H.,  
Lozada,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhinney,E., Mcleod,M.P., Meador,M., Mel,G., Metzger,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogun,M., Okwonu,G.,  
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,  
Scherer,S., Scott,G., Shen,H., Shoshari,N., Sisson,I.,  
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Umsani,K., Vasquez,L., Vera,V., Villalón,D., Vltson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlingson,S.,  
Williams,G., Williamson,A., Wlarczyk,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Welnstock,G. and Gibbs,R.

## TITLE

Direct Submission

## JOURNAL

Unpublished

## AUTHORS

2 (bases 1 to 95180)  
Morley,K.C.

## JOURNAL

Direct Submission  
Submitted (03-FEB-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

## REFERENCE

3 (bases 1 to 95180)  
Morley,K.C.

## AUTHORS

Direct Submission

## TITLE

Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One

## JOURNAL

Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One

## COMMENT

Baylor Plaza, Houston, TX 77030, USA  
 On Jul 12, 2002 this sequence version replaced g1:18846778.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 ----- Project Information  
 Center project name: GPVG  
 Center clone name: CH230-126F19  
 ----- Summary Statistics  
 Sequencing vector: plasmid  
 Chemistry: Dye-terminator Big Dye: 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 59597 bases at least Q40  
 Consensus quality: 63979 bases at least Q30  
 Consensus quality: 67799 bases at least Q20  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 48 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 1 1114: contig of 1114 bp in length  
 \* 1115 1214: gap of unknown length  
 \* 1215 2810: contig of 1596 bp in length  
 \* 2811 2910: gap of unknown length  
 \* 2911 3928: contig of 1018 bp in length  
 \* 3929 4028: gap of unknown length  
 \* 4029 5148: contig of 1120 bp in length  
 \* 5149 5248: gap of unknown length  
 \* 5249 6848: contig of 1600 bp in length  
 \* 6849 6949: gap of unknown length  
 \* 6949 8234: contig of 1306 bp in length  
 \* 8235 8355: gap of unknown length  
 \* 8355 9642: contig of 1288 bp in length  
 \* 9643 9742: gap of unknown length  
 \* 9743 11399: contig of 1657 bp in length  
 \* 11400 11499: gap of unknown length  
 \* 11500 13106: contig of 1607 bp in length  
 \* 13107 13206: gap of unknown length  
 \* 13207 15257: contig of 2051 bp in length  
 \* 15258 15358: gap of unknown length  
 \* 15358 16980: contig of 1623 bp in length  
 \* 16981 17080: gap of unknown length  
 \* 17081 18685: contig of 1605 bp in length  
 \* 18686 18785: gap of unknown length  
 \* 18786 20777: contig of 1892 bp in length  
 \* 20778 22009: gap of unknown length  
 \* 22010 22109: contig of 1232 bp in length  
 \* 22110 23662: gap of unknown length  
 \* 23663 23762: gap of unknown length  
 \* 23763 25624: contig of 1862 bp in length  
 \* 25625 25724: gap of unknown length  
 \* 25725 26730: contig of 1006 bp in length  
 \* 26731 26830: gap of unknown length  
 \* 26831 28357: contig of 1527 bp in length  
 \* 28358 28457: gap of unknown length  
 \* 28458 30373: contig of 1916 bp in length  
 \* 30374 30473: gap of unknown length  
 \* 30474 31986: contig of 1513 bp in length  
 \* 31987 32086: gap of unknown length  
 \* 32087 33490: contig of 1404 bp in length  
 \* 33491 33590: gap of unknown length  
 \* 33591 34980: contig of 1390 bp in length  
 \* 34981 35080: gap of unknown length  
 \* 35081 36720: contig of 1640 bp in length

\* 36721 36820: gap of unknown length  
 \* 36821 38004: contig of 1184 bp in length  
 \* 38005 38104: gap of unknown length  
 \* 38105 39550: contig of 1446 bp in length  
 \* 39551 39650: gap of unknown length  
 \* 39651 40651: contig of 1001 bp in length  
 \* 40652 40731: gap of unknown length  
 \* 40732 43081: contig of 2330 bp in length  
 \* 43082 43181: gap of unknown length  
 \* 43182 44469: contig of 1288 bp in length  
 \* 44470 44569: gap of unknown length  
 \* 44570 45876: contig of 1307 bp in length  
 \* 45877 45976: gap of unknown length  
 \* 45977 48240: contig of 2264 bp in length  
 \* 48241 48340: gap of unknown length  
 \* 48341 50121: contig of 1781 bp in length  
 \* 50122 50221: gap of unknown length  
 \* 50222 51854: contig of 1632 bp in length  
 \* 51855 51953: gap of unknown length  
 \* 51954 53849: contig of 1896 bp in length  
 \* 53850 53949: gap of unknown length  
 \* 53950 56898: contig of 2949 bp in length  
 \* 56899 56998: gap of unknown length  
 \* 56999 58622: contig of 1624 bp in length  
 \* 58623 58722: gap of unknown length  
 \* 58723 60532: contig of 1810 bp in length  
 \* 60533 60632: gap of unknown length  
 \* 60633 61891: contig of 1259 bp in length  
 \* 61892 61991: gap of unknown length  
 \* 61992 64149: contig of 2158 bp in length  
 \* 64150 64250: gap of unknown length  
 \* 64251 66117: contig of 1868 bp in length  
 \* 66118 66217: gap of unknown length  
 \* 66219 67638: contig of 1421 bp in length  
 \* 67639 67738: gap of unknown length  
 \* 67739 70700: contig of 2962 bp in length  
 \* 70701 70800: gap of unknown length  
 \* 70801 73188: contig of 2388 bp in length  
 \* 73189 73288: gap of unknown length  
 \* 73289 75659: contig of 2371 bp in length  
 \* 75660 75759: gap of unknown length  
 \* 75760 78787: contig of 3028 bp in length  
 \* 78788 78887: gap of unknown length  
 \* 78888 82264: contig of 3377 bp in length  
 \* 82265 82364: gap of unknown length  
 \* 82365 84919: contig of 2555 bp in length  
 \* 84920 85019: gap of unknown length  
 \* 85020 90450: contig of 5431 bp in length  
 \* 90451 90550: gap of unknown length  
 \* 90551 95180: contig of 4630 bp in length.  
 -----  
 FEATURES  
 source Location/Qualifiers  
 1. 95180  
 /organism="Rattus norvegicus"  
 /db\_xref="taxon:10116"  
 /clone="CH230-126F19"  
 -----  
 BASE COUNT 27139 a 18257 c 18189 g 25171 t 6424 others  
 ORIGIN  
 Query Match 92.0%; Score 18.4; DB 1; Length 95180;  
 Best Local Similarity 95.0%; Pred. No. 18;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 TTACTGTGTTAGTTTACTT 20  
 Db 53187 TTCTGTTTACTTATTAATT 53168  
 -----  
 RESULT 7  
 AC106466/c AC106466 106651 bp DNA linear HTG 13-JUL-2002  
 LOCUS Rattus norvegicus clone CH230-211B18, \*\*\* SEQUENCING IN PROGRESS  
 DEFINITION \*\*\* 45 unordered pieces.  
 ACCESSION AC106466



VERSION AC106466.2 GI:21735240  
 KEYWORDS HTG: HTGS\_PHASE1.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 106651)  
 Muzny,D.M., Adams,C., Adio-Odunola,B., Alt-osman,F.R., Allen,C.,  
 Alspbrooks,S.L., Amarantunga,H.C., Are,J.R., Ayale,M., Banks,T.,  
 Barbarta,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,  
 Bouck,J., Bowle,S., Bivale,M., Brown,E., Brown,M., Bryant,N.P.,  
 Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
 Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
 Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,  
 Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,  
 Homsi,F., Howard,S., Huber,J., Huiyk,S., Hume,J., Jackson,L.E.,  
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
 Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
 Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,  
 Lozado,R.D., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
 Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,S.,  
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,N.,  
 Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogun,M., Okunolu,G.,  
 Orangun,N., Oviado,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
 Rivers,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,  
 Scherer,S., Scott,G., Shen,H., Shooshari,N., Sisson,I.,  
 Sodergren,E., Sotaike,T., Sparks,A., Stanley,H., Stone,I.,  
 Sutton,A., Svatik,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
 Umanil,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,  
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,  
 Williams,G., Williamson,A., Wleczyk,R., Woodson,S., Worley,K.,  
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 Weinstein,G. and Gibbs,R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 106651)  
 Worley K.C.  
 Direct Submission  
 Submitted (12-JAN-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 106651)  
 Worley K.C.  
 Direct Submission  
 Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Jul 12, 2002 this sequence version replaced gi:18138988.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GKYR  
 Center clone name: CH230-211B18  
 ----- Summary Statistics  
 Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 79275 bases at least Q40  
 Consensus quality: 83497 bases at least Q30  
 Consensus quality: 87549 bases at least Q20  
 -----

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 45 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 1130: contig of 1130 bp in length  
 1131 1230: gap of unknown length  
 1231 2243: contig of 1013 bp in length  
 2244 2343: gap of unknown length  
 2344 3369: contig of 1026 bp in length  
 3370 3469: gap of unknown length  
 3470 5003: contig of 1534 bp in length  
 5004 5103: gap of unknown length  
 5104 6657: contig of 1534 bp in length  
 6658 6757: gap of unknown length  
 6758 7840: contig of 1083 bp in length  
 7841 7940: gap of unknown length  
 7941 10020: contig of 2080 bp in length  
 10021 10120: gap of unknown length  
 10121 11477: contig of 1357 bp in length  
 11478 11577: gap of unknown length  
 11578 13365: contig of 1788 bp in length  
 13366 13465: gap of unknown length  
 13466 14853: contig of 1388 bp in length  
 14854 14953: gap of unknown length  
 14954 16290: contig of 1337 bp in length  
 16291 16390: gap of unknown length  
 16391 17509: contig of 1119 bp in length  
 17510 17609: gap of unknown length  
 17610 19000: contig of 1391 bp in length  
 19001 19100: gap of unknown length  
 19101 20944: contig of 1844 bp in length  
 20945 22651: gap of unknown length  
 22652 22751: contig of 1607 bp in length  
 22752 25058: contig of 2307 bp in length  
 25059 25158: gap of unknown length  
 25159 26816: contig of 1658 bp in length  
 26817 26916: gap of unknown length  
 26917 28641: contig of 1725 bp in length  
 28642 28741: gap of unknown length  
 28742 30799: contig of 2038 bp in length  
 30800 30899: gap of unknown length  
 30900 32351: contig of 1452 bp in length  
 32352 32451: gap of unknown length  
 32452 34522: contig of 2071 bp in length  
 34523 34622: gap of unknown length  
 34623 36509: contig of 1887 bp in length  
 36510 36609: gap of unknown length  
 36610 38251: contig of 1642 bp in length  
 38252 38351: gap of unknown length  
 38352 40309: contig of 1956 bp in length  
 40310 40409: gap of unknown length  
 40410 42792: contig of 2383 bp in length  
 42793 42892: gap of unknown length  
 42893 45780: contig of 2888 bp in length  
 45781 45880: gap of unknown length  
 45881 49321: contig of 3341 bp in length  
 49322 49321: gap of unknown length  
 49322 51468: contig of 2147 bp in length  
 51469 51568: gap of unknown length  
 51569 53749: contig of 2181 bp in length

53750 53849: gap of unknown length  
 53850 55748: contig of 1899 bp in length  
 55849 55748: gap of unknown length  
 55849 59788: contig of 3940 bp in length  
 59789 59888: gap of unknown length  
 59889 62648: contig of 2761 bp in length  
 62650 62749: gap of unknown length  
 62750 65930: contig of 3181 bp in length  
 65931 66030: gap of unknown length  
 66031 68484: contig of 2454 bp in length  
 68485 68584: gap of unknown length  
 68585 70509: contig of 1925 bp in length  
 70510 70609: gap of unknown length  
 70610 73606: contig of 2997 bp in length  
 73607 76448: contig of 2742 bp in length  
 76449 76548: gap of unknown length  
 76549 79392: contig of 2844 bp in length  
 79393 79492: gap of unknown length  
 79493 82319: contig of 2827 bp in length  
 82320 82419: gap of unknown length  
 82420 85792: contig of 3373 bp in length  
 85793 85892: gap of unknown length  
 85893 88434: contig of 2541 bp in length  
 88434 88533: gap of unknown length  
 88534 92245: contig of 3712 bp in length  
 92246 92345: gap of unknown length  
 92346 96997: contig of 4652 bp in length  
 96998 97097: gap of unknown length  
 97098 102010: contig of 4913 bp in length  
 102011 102110: gap of unknown length  
 102111 106651: contig of 4541 bp in length.  
 Location/Qualifiers  
 1. 106651  
 /organism="Rattus norvegicus"  
 /db\_xref="taxon:10116"  
 /clone="CH230-211B18"

BASE COUNT 29303 a 19724 c 19632 g 30261 t 7731 others  
 ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 106651;  
 Best Local Similarity 95.0%; Pred. No. 19;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTACTGTCTAGTATTAGTT 20  
 ||||| ||||| ||||| |||||  
 Db 83573 TTACTGTCTAGTATTAGTT 83554

RESULT 8  
 AC128956  
 LOCUS  
 DEFINITION  
 Rattus norvegicus clone CH230-500P2, \*\*\* SEQUENCING IN PROGRESS  
 \*\*\* 38 unordered pieces.

AC128956 150326 bp DNA linear HTG 24-JUL-2002  
 AC128956.1 GI:21953705  
 HTG: HTGS\_PHASE1.  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 150326)

REFERENCE  
 AUTHORS  
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
 Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,  
 Barberia,J., Benton,J., Bimge,K., Blankenburg,K., Bonnin,D.,  
 Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,  
 Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
 Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

COMMENT

Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
 Earnhart,C., Edgar,D., Edwards,C.C., Elhai,C., Escoto,M.,  
 Fall,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
 Gabril,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
 Gorrill,J.H., Guenara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
 Harris,C., Harris,K., Hart,M., Havlak,P., Hayes,A., Hernandez,J.,  
 Hernandez,O., Hodgeson,A., Hognes,M., Holloway,C., Hollins,B.,  
 Homsi,F., Howard,S., Huber,J., Huik,S., Hume,J., Jackson,L.E.,  
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
 Karlsson,E., Kelly,S., Khan,U., Kling,L., Korah,J., Kovar,C.,  
 Kratovic,J., Kureishi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
 Li,J., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,W., Louisege,H.,  
 Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
 Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,  
 Miner,G., Miner,Z., Mitchell,T., Mohabath,K., Morgan,M., Morris,S.,  
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
 Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okunolu,G.,  
 Otangunye,N., Oyedokun,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
 Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,  
 Scherer,S., Scott,G., Shen,H., Shooshbari,N., Sisson,I.,  
 Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,  
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
 Tansey,J., Taylor,C., Taylor,T., Teifrod,B., Thomas,N., Thomas,S.,  
 Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,  
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,  
 Williams,G., Williamson,A., Wlezyk,R., Wooden,S., Worley,K.,  
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 Weinstein,G., and Gibbs,R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 150326)  
 Morley,K.C.  
 Direct Submission  
 Submitted (24-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 -----  
 Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: http://www.hgsc.bcm.tmc.edu/  
 Contact: hgsc-help@bcm.tmc.edu  
 -----  
 Project Information  
 Center project name: KBPH  
 Center clone name: CH230-500P2  
 -----  
 Summary Statistics  
 Sequencing vector: Plasmid:  
 Chemistry: Dye-terminator Big Dye: 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 110267 bases at least Q40  
 Consensus quality: 116757 bases at least Q30  
 Consensus quality: 121521 bases at least Q20  
 -----  
 NOTE: Estimated insert size may differ from sequence length  
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_data.html).  
 NOTE: This is a 'working draft' sequence. It currently  
 consists of 38 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.  
 1 1047: contig of 1047 bp in length  
 \* 1048 1147: gap of unknown length  
 \* 1148 2218: contig of 1071 bp in length  
 \* 2219 2318: gap of unknown length  
 \* 2319 3360: contig of 1042 bp in length  
 \* 3361 3460: gap of unknown length  
 \* 3461 5074: contig of 1614 bp in length  
 \* 5075 5174: gap of unknown length  
 \* 5175 6294: contig of 1120 bp in length

```

* 6295 6394: gap of unknown length
* 6395 8169: contig of 1775 bp in length
* 8170 8269: gap of unknown length
* 8270 9298: contig of 1029 bp in length
* 9299 9398: gap of unknown length
* 9399 10923: contig of 1525 bp in length
* 10924 11023: gap of unknown length
* 11024 12507: contig of 1484 bp in length
* 12508 12607: gap of unknown length
* 12608 14841: contig of 2234 bp in length
* 14842 14941: gap of unknown length
* 14942 17290: contig of 2349 bp in length
* 17291 17390: gap of unknown length
* 17391 18790: contig of 1400 bp in length
* 18791 18890: gap of unknown length
* 18891 20400: contig of 1510 bp in length
* 20401 20500: gap of unknown length
* 20501 22551: contig of 2051 bp in length
* 22552 22651: gap of unknown length
* 22652 24396: contig of 1745 bp in length
* 24397 24496: gap of unknown length
* 24497 26180: contig of 1684 bp in length
* 26181 26280: gap of unknown length
* 26281 28087: contig of 1807 bp in length
* 28088 28187: gap of unknown length
* 28188 30547: contig of 2360 bp in length
* 30548 30647: gap of unknown length
* 30648 33337: contig of 2690 bp in length
* 33338 33437: gap of unknown length
* 33438 36007: contig of 2570 bp in length
* 36008 36107: gap of unknown length
* 36108 39900: contig of 3793 bp in length
* 39901 40000: gap of unknown length
* 40001 43740: contig of 3740 bp in length
* 43741 43840: gap of unknown length
* 43841 47274: contig of 3434 bp in length
* 47275 47374: gap of unknown length
* 47375 52052: contig of 4678 bp in length
* 52053 52152: gap of unknown length
* 52153 55896: contig of 3744 bp in length
* 55897 55996: gap of unknown length
* 55997 61630: contig of 5634 bp in length
* 61631 61730: gap of unknown length
* 61731 65336: contig of 3606 bp in length
* 65337 65436: gap of unknown length
* 65437 69824: contig of 4388 bp in length
* 69825 69924: gap of unknown length
* 69925 74153: contig of 4229 bp in length
* 74154 74253: gap of unknown length
* 74254 80624: contig of 6371 bp in length
* 80625 85103: gap of unknown length
* 85104 85203: contig of 4379 bp in length
* 85204 90043: gap of unknown length
* 90044 90143: gap of unknown length
* 90144 94288: contig of 4145 bp in length
* 94289 94388: gap of unknown length
* 94389 99899: contig of 5511 bp in length
* 99900 107750: gap of unknown length
* 100000 107751: contig of 7751 bp in length
* 107751 107850: gap of unknown length
* 114038 114037: contig of 6187 bp in length
* 114038 114137: gap of unknown length
* 114138 132540: contig of 18403 bp in length
* 132541 132641: gap of unknown length
* 132641 130326: contig of 17686 bp in length.

```

## FEATURES

Location/Qualifiers

1..150326

/organism="Rattus norvegicus"

/db\_xref="taxon:10116"

/clone="CH230-500b2"

## BASE COUNT

36075 a 34635 c 33943 g 36814 t 8859 others

## ORIGIN

```

Query Match          92.0%; Score 18.4; DB 1; Length 150326;
Best Local Similarity 95.0%; Pred. No. 19;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

> 1 TTAGTTGTTAGTTATTAGTT 20
      |||
Db 114899 TTAGTTGTTAGTTGTTAGTT 114918

```

## RESULT 9

AC129017/c

LOCUS

DEFINITION

AC129017

AC129017.1 GI:21954045

KEYWORDS

HTG; HTGS\_PHASE1; HTGS\_DRAFT.

SOURCE

MUS MUSCULUS

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

```

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watson.wustl.edu
Project Information
Center project name: M.BB041424

----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 148742 bases at least Q40
Consensus quality: 149676 bases at least Q30
Consensus quality: 150242 bases at least Q20
Insert size: 152557; sum-of-ctgts
Quality coverage: 6.72 in Q20 bases; agarose-fp
Quality coverage: 6.60 in Q20 bases; sum-of-ctgts

* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

```

1 1133: contig of 1133 bp in length
* 1134 1233: gap of unknown length
* 1234 4269: contig of 3036 bp in length
* 4270 4369: gap of unknown length
* 4370 8447: contig of 4078 bp in length
* 8448 8547: gap of unknown length
* 8548 20096: contig of 11449 bp in length
* 20097 20196: gap of unknown length
* 20197 40259: contig of 20063 bp in length
* 40260 40359: gap of unknown length
* 40360 63542: contig of 23183 bp in length

```

```

* 63543 63642: gap of unknown length
* 63643 85165: contig of 21523 bp in length
* 85166 85265: gap of unknown length
* 85266 118186: contig of 32921 bp in length
* 118187 118286: gap of unknown length
* 118287 152324: contig of 34038 bp in length.
FEATURES
Source
1.152324
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="X"
/clone="RP24-41A24"
1.1133
/note="assembly_name:Contig13"
1234.4269
/note="assembly_name:Contig14"
4370.8447
/note="assembly_name:Contig15
clone_end:sp6
vector_side:left"
8348.20096
/note="assembly_name:Contig16"
20197.40259
/note="assembly_name:Contig17"
40360.63542
/note="assembly_name:Contig18"
63643.85165
/note="assembly_name:Contig19
clone_end:T7
vector_side:left"
85266.118186
/note="assembly_name:Contig20"
118287.152324
/note="assembly_name:Contig21"
BASE COUNT 49091 a 27882 c 28512 g 45991 t 848 others
ORIGIN
Query Match 92.0% Score 18.4; DB 1; Length 152324;
Best Local Similarity 95.0% Pred. No. 19;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TTAGTGTAGTATTAGT 20
|||||
Db 28012 TTAGTGTAGTATTAGT 27993
RESULT 10
AC11380 164552 bp DNA linear HTG 13-JUN-2002
LOCUS Rattus norvegicus clone CH230-130H20, *** SEQUENCING IN PROGRESS
DEFINITION
***, 58 unordered pieces.
ACCESSION AC11380
VERSION AC11380.2 GI:21735908
KEYWORDS HTG: HTGS PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 164552)
Muzny,D.M., Adams,C., Adio-oduola,B., All-osman,F.R., Allen,C.,
Alsdbrook,S.L., Amaralunge,H.C., Aye,J.R., Ayele,M., Banks,T.,
Barbarta,J., Benton,J., Bimaye,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowler,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carton,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douglatte,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,

```

```

Gabriel,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homs,F., Howard,S., Huber,J., Hulik,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolyet,S., Joudah,S.,
Karlisson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kralovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,W., Louised,H.,
Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Merindade,A., Martinez,E.,
Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Mosser,M., Neal,D., Newton,N., Newton,S., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenko,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Severy,G.,
Scherer,S., Scott,G., Shen,H., Shoochari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,Y., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G., and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 164552)
Morley,K.C.
JOURNAL
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 164552)
Morley,K.C.
JOURNAL
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 12, 2002 this sequence version replaced gi:18701144.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GMDJ
Center clone name: CH230-130H20
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 105093 bases at least Q40
Consensus quality: 110397 bases at least Q30
Consensus quality: 114114 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 58 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1002: contig of 1002 bp in length
* 1003 1102: gap of unknown length
* 1103 2301: contig of 1199 bp in length
* 2302 2401: gap of unknown length
* 2402 3505: contig of 1104 bp in length

```

```

* 3506 3605: gap of unknown length
* 3606 4783: contig of 1178 bp in length
* 4784 4883: gap of unknown length
* 4884 6056: contig of 1173 bp in length
* 6057 6156: gap of unknown length
* 6157 7495: contig of 1339 bp in length
* 7496 7595: gap of unknown length
* 7596 8658: contig of 1063 bp in length
* 8659 8758: gap of unknown length
* 8759 10238: contig of 1480 bp in length
* 10239 10338: gap of unknown length
* 10339 12220: contig of 1882 bp in length
* 12221 12320: gap of unknown length
* 12321 13477: contig of 1157 bp in length
* 13478 13577: gap of unknown length
* 13578 14605: contig of 1028 bp in length
* 14606 14705: gap of unknown length
* 14706 15755: contig of 1050 bp in length
* 15756 15855: gap of unknown length
* 15856 17702: contig of 1847 bp in length
* 17703 17802: gap of unknown length
* 17803 19472: contig of 1670 bp in length
* 19473 19572: gap of unknown length
* 19573 21485: contig of 1913 bp in length
* 21486 21585: gap of unknown length
* 21586 23432: contig of 1847 bp in length
* 23433 23532: gap of unknown length
* 23533 25029: contig of 1497 bp in length
* 25030 25129: gap of unknown length
* 25130 26659: contig of 1530 bp in length
* 26660 26759: gap of unknown length
* 26760 28350: contig of 1591 bp in length
* 28351 28450: gap of unknown length
* 28451 30077: contig of 1627 bp in length
* 30078 30177: gap of unknown length
* 30178 31793: contig of 1616 bp in length
* 31794 31893: gap of unknown length
* 31894 33673: contig of 1780 bp in length
* 33674 33773: gap of unknown length
* 33774 34989: contig of 1216 bp in length
* 34990 35089: gap of unknown length
* 35090 36437: contig of 1348 bp in length
* 36438 36537: gap of unknown length
* 36538 37948: contig of 1411 bp in length
* 37949 38048: gap of unknown length
* 38049 39244: contig of 1196 bp in length
* 39245 39344: gap of unknown length
* 39345 41409: contig of 2065 bp in length
* 41410 41509: gap of unknown length
* 41510 43089: contig of 1580 bp in length
* 43090 43189: gap of unknown length
* 43190 45987: contig of 2798 bp in length
* 45988 46087: gap of unknown length
* 46088 47997: contig of 1910 bp in length
* 47998 48097: gap of unknown length
* 48098 50569: contig of 2472 bp in length
* 50570 50669: gap of unknown length
* 50670 53411: contig of 2742 bp in length
* 53412 53511: gap of unknown length
* 53512 56398: contig of 2887 bp in length
* 56399 56498: gap of unknown length
* 56499 58805: contig of 2307 bp in length
* 58806 58905: gap of unknown length
* 58906 61202: contig of 2297 bp in length
* 61203 61302: gap of unknown length
* 61303 63871: contig of 2569 bp in length
* 63872 63971: gap of unknown length
* 63972 66802: contig of 2831 bp in length
* 66803 66902: gap of unknown length
* 66903 69598: contig of 2696 bp in length
* 69599 72577: gap of unknown length
* 72578 72677: contig of 2879 bp in length

```

```

Query Match          92.0%; Score 18.4; DB 1; Length 164552;
Best Local Similarity 95.0%; Pred. No. 19;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 TTAGTGTAGTTATTAGTT 20

```

```

Db 130778 TTAGTGTAGTTAGTTAGTT 130797

```

```

RESULT 11
LOCUS          AC113369          176400 bp      DNA      linear      HTG 01-MAR-2002
DEFINITION    Homo sapiens chromosome 5 clone RP11-119p10, WORKING DRAFT
ACCESSION     AC113369
VERSION       AC113369.1  GI:19033483
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVERIN.
SOURCE        Homo sapiens.
ORGANISM      Homo sapiens
REFERENCE     1 (bases 1 to 176400)
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
JOURNAL       Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     2 (bases 1 to 176400)
AUTHORS      Unpublished
TITLE         DOE Joint Genome Institute.
JOURNAL       DOE Joint Genome Institute.
COMMENT       Submitted (01-MAR-2002) Production Sequencing Facility, DOE Joint
              Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
              -----Genome Center
              Center: Joint Genome Institute
              Center Code: JGI
              Web site: http://www.jgi.doe.gov
              -----
              Project Information
              Center Project Name: 452431
              Center Clone name: RPCI-11_119p10
              -----
              Summary Statistics
              Consensus quality: 175634 bases at least Q40
              Consensus quality: 176221 bases at least Q30
              Consensus quality: 176375 bases at least Q20

```

```

* Estimated insert size: 182000; agarose-fp estimation
* Estimated insert size: 176400; sum-of-ctnigs estimation
* Quality coverage: 12.49 in Q20 bases; agarose-fp estimation
* Quality coverage: 12.88 in Q20 bases; sum-of-ctnigs estimation
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 ctnigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the ctnigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 176400: contig of 176400 bp in length.

```

FEATURES	Location/Qualifiers
source	1. .176400

BASE COUNT	55190	a	37584	c	34956	g	48670	t
ORIGIN								

Query Match	92.0%;	Score 18.4;	DB 1;	Length 176400;
Best Local Similarity	95.0%;	Pred. No. 20;		
Matches 19;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0

Qy	1	TTAGTGTGTTAGTATTAGCTT	20
Db	78311	TTAGTATTAGTATTAGCTT	78330

RESULT	12
AC115176/c	
LOCUS	AC115176
DEFINITION	Rattus norvegicus clone CH230-138I22, *** SEQUENCING IN PROGRESS
ACCSSION	AC115176
VERSION	AC115176.5 GI:21738734
KEYWORDS	HTG; HTGS_PHASE1.
SOURCE	Norway rat.
ORGANISM	Rattus norvegicus

TITLE	Direct Submission
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 178712)
AUTHORS	Worley,K.C.
TITLE	Direct Submission
JOURNAL	Submitted (15-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	3 (bases 1 to 178712)
AUTHORS	Worley,K.C.
TITLE	Direct Submission
JOURNAL	Submitted (13-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT	On Jul 12, 2002 this sequence version replaced g1:20467608.

*	1	1025: contig of 1025 bp in length
*	1026	1125: gap of unknown length
*	1126	2503: contig of 1378 bp in length
*	2504	2603: gap of unknown length
*	2604	3547: contig of 1044 bp in length
*	3548	3747: gap of unknown length
*	3748	4848: contig of 1101 bp in length
*	4849	4548: gap of unknown length
*	4849	6516: contig of 1568 bp in length
*	6517	6616: gap of unknown length
*	6617	7633: contig of 1017 bp in length
*	7634	7733: gap of unknown length
*	7734	9412: contig of 1679 bp in length
*	9413	9512: gap of unknown length
*	9513	10664: contig of 1052 bp in length
*	10565	10664: gap of unknown length
*	10665	11733: contig of 1069 bp in length
*	11734	11733: gap of unknown length
*	11834	13300: contig of 1467 bp in length

13301 13400: gap of unknown length  
13401 15196: contig of 1796 bp in length  
15197 15296: gap of unknown length  
15297 16682: contig of 1386 bp in length  
16683 16782: gap of unknown length  
16783 17945: contig of 1163 bp in length  
17946 18045: gap of unknown length  
18046 19776: contig of 1731 bp in length  
19777 19876: gap of unknown length  
19877 21957: contig of 2081 bp in length  
21958 22057: gap of unknown length  
22058 23188: contig of 1131 bp in length  
23189 23288: gap of unknown length  
23289 24464: contig of 1176 bp in length  
24465 24564: gap of unknown length  
24565 25908: contig of 1344 bp in length  
25909 26008: gap of unknown length  
26009 27814: contig of 1806 bp in length  
27815 27914: gap of unknown length  
27915 29387: contig of 1473 bp in length  
29388 29487: gap of unknown length  
29488 30801: contig of 1314 bp in length  
30802 30901: gap of unknown length  
30902 32199: contig of 1298 bp in length  
32200 32299: gap of unknown length  
32300 33740: contig of 1441 bp in length  
33741 33840: gap of unknown length  
33841 34992: contig of 1152 bp in length  
34993 35092: gap of unknown length  
35093 37835: contig of 2743 bp in length  
37836 37935: gap of unknown length  
37936 41035: contig of 3100 bp in length  
41036 41135: gap of unknown length  
41136 43185: contig of 2050 bp in length  
43186 43285: gap of unknown length  
43286 45335: contig of 2050 bp in length  
45336 45435: gap of unknown length  
45436 46618: contig of 1183 bp in length  
46619 46719: gap of unknown length  
46719 49009: contig of 2291 bp in length  
49010 49109: gap of unknown length  
49110 52162: contig of 3053 bp in length  
52163 52262: gap of unknown length  
52263 54282: contig of 2020 bp in length  
54283 54382: gap of unknown length  
54383 56291: contig of 1909 bp in length  
56292 56391: gap of unknown length  
56392 58037: contig of 1646 bp in length  
58038 58137: gap of unknown length  
58138 62042: contig of 3905 bp in length  
62043 62142: gap of unknown length  
62143 64769: contig of 2627 bp in length  
64770 64869: gap of unknown length  
64870 67320: contig of 2451 bp in length  
67321 67420: gap of unknown length  
67421 70041: contig of 2621 bp in length  
70042 70141: gap of unknown length  
70142 73105: contig of 2964 bp in length  
73106 73205: gap of unknown length  
73206 76419: contig of 3214 bp in length  
76420 76519: gap of unknown length  
76520 79593: contig of 3074 bp in length  
79594 79693: gap of unknown length  
79694 82683: contig of 2990 bp in length  
82684 82783: gap of unknown length  
82784 85451: contig of 2668 bp in length  
85452 85551: gap of unknown length  
85552 89457: contig of 3906 bp in length  
89458 89557: gap of unknown length  
89558 91719: contig of 2162 bp in length  
91720 91819: gap of unknown length  
91820 95602: contig of 3783 bp in length  
95603 95702: gap of unknown length

Query Match 92.0%; Score 18.4; DB 1; Length 178712;  
Best Local Similarity 95.0%; Pred. No. 20;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TTAGTGTAGTATTAGTT 20  
Db 109564 TTAGTGTAGTGTAGTT 109545

RESULT 13  
AC012589  
LOCUS  
DEFINITION Homo sapiens chromosome 5 clone RP11-119P10 map 5, WORKING DRAFT  
SEQUENCE, 23 unordered pieces.  
ACCESSION AC012589  
VERSION AC012589.5 GI:7387335  
KEYWORDS HTG: HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Eukaryota; Euthera; Primates; Carnivora; Mammalia; Homo.  
1 (bases 1 to 178970)  
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,  
Baldwin, J., Barna, N., Beckert, R., Boguslavsky, L., Boukhalter, B.,  
Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,  
Cooke, P., Dearrellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,  
Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,  
Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howland, J., Johnson, R., Jones, C., Kahn, L., Karatas, A., Klein, J.,  
Lehoczky, J., Lieu, C., Locke, K., MacDonald, P., Marquis, N.,  
McEwan, P., McGuck, A., McKernan, K., McLaughlin, J., Melidm, J.,  
Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,  
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Tefaye, S., Tiller, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,  
Wymann, D., Ye, W. J., Zimmer, A. and Zody, M.

TITLE Direct Submission  
JOURNAL Submitted (30-Oct-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 178970)  
REFERENCE  
AUTHORS

Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A.,  
Camarata, J., Campopiano, A., Chang, J., Choepel, Y., Colangelo, M.,  
Collins, S., Collymore, A., Cooke, P., Dearrellano, K., Dewar, K.,  
Diaz, J., S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D.,  
Galagan, J., Gardyna, S., Glnde, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Huime, N.,  
Iliev, I., Johnson, R., Jones, C., Karatas, A., Labocque, K.,  
Lamazares, R., Landers, T., Lehoczky, J., Levine, R., Liu, G.,  
Maclean, C., MacDonald, P., Marquis, N., Matthews, C., McCarthy, M.,  
McEwan, P., McKernan, K., McPheters, R., Melidm, J., Melidm, J.,  
Milhova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C.,  
Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J.,

TITLE  
JOURNAL  
COMMENT

Peterson, K., Phunhkhang, P., Pierre, N., Pollara, V., Raymond, C.,  
Retta, R., Rieback, M., Riley, R., Riese, C., Rogov, P., Roman, D.,  
Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S.,  
Severy, P., Sounez, C., Spencer, B., Stange-Thomann, N.,  
Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S.,  
Theodore, J., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Apr 3, 2000 this sequence version replaced g1:6910762.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

## Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: <http://www.seq.wi.mit.edu>Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

Project Information

Center project name: L3738

Center clone name: 119.P.10

## Summary Statistics

Sequencing vector: M13; M7815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 160436 bases at least Q40  
Consensus quality: 168973 bases at least Q30  
Consensus quality: 173202 bases at least Q20  
Insert size: 177000; agarose-fp  
Insert size: 176770; sum-of-ctigs  
Quality coverage: 4.0 in Q20 bases; sum-of-ctigs  
Quality coverage: 4.0 in Q20 bases; sum-of-ctigs

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 23 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1  
\* 1012 1111: contig of 1011 bp in length  
\* 1112 2722: contig of 1611 bp in length  
\* 2723 2822: gap of 100 bp  
\* 2823 4300: contig of 1478 bp in length  
\* 4401 4400: gap of 100 bp  
\* 4401 5425: contig of 1025 bp in length  
\* 5426 5525: gap of 100 bp  
\* 5526 7534: contig of 2009 bp in length  
\* 7535 7634: gap of 100 bp  
\* 7635 9780: contig of 2146 bp in length  
\* 9781 9880: gap of 100 bp  
\* 9881 12462: contig of 2582 bp in length  
\* 12463 12562: gap of 100 bp  
\* 12563 15626: contig of 3064 bp in length  
\* 15627 15726: gap of 100 bp  
\* 15727 17635: contig of 1909 bp in length  
\* 17636 17735: gap of 100 bp  
\* 17736 19894: contig of 2159 bp in length  
\* 19895 19994: gap of 100 bp  
\* 19995 25298: contig of 5304 bp in length  
\* 25299 25398: gap of 100 bp  
\* 25399 30406: contig of 5008 bp in length  
\* 30407 30506: gap of 100 bp  
\* 30507 38666: contig of 8160 bp in length  
\* 38667 38766: gap of 100 bp  
\* 38767 45449: contig of 6683 bp in length  
\* 45450 45549: gap of 100 bp  
\* 45550 55180: contig of 9631 bp in length  
\* 55181 55280: gap of 100 bp

## FEATURES

source

\* 55281 63752: contig of 8472 bp in length  
\* 63753 63852: gap of 100 bp  
\* 63853 75295: contig of 11443 bp in length  
\* 75296 75395: gap of 100 bp  
\* 75396 87583: contig of 12188 bp in length  
\* 87584 87683: gap of 100 bp  
\* 87684 98245: contig of 10562 bp in length  
\* 98246 98345: gap of 100 bp  
\* 98346 111685: contig of 13340 bp in length  
\* 111686 111785: gap of 100 bp  
\* 111786 131002: contig of 19217 bp in length  
\* 131003 131102: gap of 100 bp  
\* 131103 131207: contig of 20968 bp in length  
\* 131207 152170: gap of 100 bp  
\* 152171 178970: contig of 26800 bp in length.  
\* 152171 178970: Location/Qualifiers  
1. 178970  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/map="5"  
/clone="RP11-119P10"  
/clone\_lib="RP11 Human Male BAC"  
1. 1011  
/note="assembly-fragment"  
1112. 2722  
/note="assembly-fragment"  
2823. 4300  
/note="assembly-fragment"  
4401. 5425  
/note="assembly-fragment"  
5526. 7534  
/note="assembly-fragment"  
7635. 9780  
/note="assembly-fragment"  
9881. 12462  
/note="assembly-fragment"  
12563. 15626  
/note="assembly-fragment"  
15727. 17635  
/note="assembly-fragment"  
17736. 19894  
/note="assembly-fragment"  
19995. 25298  
/note="assembly-fragment"  
25399. 30406  
/note="assembly-fragment"  
30507. 38666  
/note="assembly-fragment"  
38767. 45449  
/note="assembly-fragment"  
45550. 55180  
/note="assembly-fragment"  
55281. 63752  
/note="assembly-fragment"  
63853. 75295  
/note="assembly-fragment"  
75396. 87583  
/note="assembly-fragment"  
87684. 98245  
/note="assembly-fragment"  
98346. 111685  
/note="assembly-fragment"  
111786. 131002  
/note="assembly-fragment"  
131103. 152070  
/note="assembly-fragment"  
clone\_end:SP6  
vector\_side:right  
152171. 178970  
/note="assembly-fragment"  
clone\_end:T7  
vector\_side:left"



```

BASE COUNT      54765 a 37234 c 35721 g 49050 t 2200 others
ORIGIN
Query Match      92.0%; Score 18.4; DB 1; Length 178970;
Best Local Similarity 95.0%; Pred. No. 20;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 1 TTAGTTGTTAGTTATTAGTT 20
      ||||| ||||| ||||| |||||
Db 22509 TTAGTTATTAGTTATTAGTT 22528

RESULT 14
AC122335      201132 bp   DNA      linear   HTG 02-AUG-2002
LOCUS      Mus musculus chromosome UNK clone RP23-350F7, WORKING DRAFT
DEFINITION  SEQUENCE, 10 unordered pieces.
ACCESSION  AC122335.2 GI:22091367
VERSION    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS   house mouse.
SOURCE     Mus musculus
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 201132)
AUTHORS   McPherson,J.D. and Waterston,R.H.
TITLE     The sequence of Mus musculus clone
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 201132)
AUTHORS   McPherson,J.D. and Waterston,R.H.
TITLE     Direct Submission
JOURNAL   Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
            3 (bases 1 to 201132)
            McPherson,J.D. and Waterston,R.H.
REFERENCE  Submitted (02-AUG-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
COMMENT    On Aug 2, 2002 this sequence version replaced gi:21105196.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watson.wustl.edu
----- Project Information -----
Center project name: M_BA0350F07
----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: Plasmid; 100%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; Version 0.990319
Consensus quality: 0 bases at least Q40
Consensus quality: 0 bases at least Q30
Consensus quality: 0 bases at least Q20
Insert size: 191000; agarose-fp
Insert size: 199573; sum-of-contigs
Quality coverage: 14.48 in Q20 bases; sum-of-contigs
Quality coverage: 11.45 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 2374: contig of 2374 bp in length
* 2375 2474: gap of unknown length
* 2475 7198: contig of 4724 bp in length

```

```

* 7199 7298: gap of unknown length
* 7299 13971: contig of 6673 bp in length
* 13972 14071: gap of unknown length
* 14072 23201: contig of 9130 bp in length
* 23202 23301: gap of unknown length
* 23302 36024: contig of 12723 bp in length
* 36025 36125: gap of unknown length
* 36125 55431: contig of 19306 bp in length
* 55431 55531: gap of unknown length
* 55531 75591: contig of 20060 bp in length
* 75591 75691: gap of unknown length
* 75691 100853: contig of 25163 bp in length
* 100854 100953: gap of unknown length
* 100954 200373: contig of 99420 bp in length
* 200374 200473: gap of unknown length
* 200474 201132: contig of 659 bp in length.
FEATURES
source
1..201132
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP23-350F7"
1..2374
/note="assembly_name:Contig11"
2475..7198
/note="assembly_name:Contig12"
7299..13971
/note="assembly_name:Contig13"
14072..23201
/note="assembly_name:Contig14"
23302..36024
/note="assembly_name:Contig15"
36125..55430
/note="assembly_name:Contig16"
55531..75590
/note="assembly_name:Contig17"
75691..100853
/note="assembly_name:Contig18"
100954..200373
/note="assembly_name:Contig19"
200474..201132
/note="assembly_name:Contig9"
BASE COUNT      56228 a 41046 c 42038 g 60912 t 908 others
ORIGIN
Query Match      92.0%; Score 18.4; DB 1; Length 201132;
Best Local Similarity 95.0%; Pred. No. 20;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 1 TTAGTTGTTAGTTATTAGTT 20
      ||||| ||||| ||||| |||||
Db 143770 TTAGTTATTAGTTATTAGTT 143789

RESULT 15
AL844536      241585 bp   DNA      linear   HTG 09-AUG-2002
LOCUS      Mus musculus chromosome 2 clone RP23-22A15, *** SEQUENCING IN
DEFINITION  PROGRESS ***, 11 unordered pieces.
ACCESSION  AL844536.3 GI:22204804
VERSION    HTG; HTGS_PHASE1.
KEYWORDS   house mouse.
SOURCE     Mus musculus
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 241585)
AUTHORS   Plumb,B.
TITLE     Direct Submission
JOURNAL   Submitted (02-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
            On Aug 11, 2002 this sequence version replaced gi:21955751.
COMMENT

```

```
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: DM22A15
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 236896 bases at least Q40
Consensus quality: 238326 bases at least Q30
Consensus quality: 239491 bases at least Q20
Insert size: 240585; sum-of-coverage
Insert size: 207539; 1.7% error; agarose-fp
Quality coverage: 4.96x in Q20 bases; sum-of-coverage
Coverage: 5.80x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 26384: contig of 26384 bp in length
* 26385 26484: gap of 100 bp
* 26485 38315: contig of 11831 bp in length
* 38316 38415: gap of 100 bp
* 38416 41944: contig of 3529 bp in length
* 41945 42044: gap of 100 bp
* 42045 56520: contig of 14476 bp in length
* 56521 56620: gap of 100 bp
* 56621 77754: contig of 21134 bp in length
* 77755 77854: gap of 100 bp
* 77855 96164: contig of 18310 bp in length
* 96165 96264: gap of 100 bp
* 96265 101917: contig of 5653 bp in length
* 101918 102017: gap of 100 bp
* 102018 146947: contig of 44930 bp in length
* 146948 147047: gap of 100 bp
* 147048 187542: contig of 40495 bp in length
* 187543 187642: gap of 100 bp
* 187643 192747: contig of 5105 bp in length
* 192748 192847: gap of 100 bp
* 192848 241585: contig of 48738 bp in length.
*
* Location/Qualifiers
*   1..241585
*     /organism="Mus musculus"
*     /db_xref="taxon:10090"
*     /chromosome="2"
*     /clone="RP23-22A15"
*     /clone_1b="RPC1-23"
*   1..26384
*     /note="assembly_fragment:01220"
*     fragment_chain:1"
*   26485..38315
*     /note="assembly_fragment:01223"
*     fragment_chain:1"
*   38416..41944
*     /note="assembly_fragment:01396"
*     fragment_chain:1"
*   42045..56520
*     /note="assembly_fragment:01233"
*     fragment_chain:1"
*   56621..77754
*     /note="assembly_fragment:00572"
*     fragment_chain:2"
*   77855..96164
*     /note="assembly_fragment:01046"
*     fragment_chain:2"
*   96265..101917
*     /note="assembly_fragment:01569"
*     fragment_chain:2"
*   102018..146947
*     /note="assembly_fragment:00752"
*     fragment_chain:2"
*   clone_end:T7
*     vector_side:right"
*   147048..187542
*     /note="assembly_fragment:00109"
*     fragment_chain:2"
*   187643..192747
*     /note="assembly_fragment:00452"
*     fragment_chain:2"
*   192848..241585
*     /note="assembly_fragment:00041"
*     fragment_chain:2"
```

```
misc_feature /note="assembly_fragment:01569"
fragment_chain:2"
misc_feature /note="assembly_fragment:00752"
fragment_chain:2"
misc_feature /note="assembly_fragment:00109"
fragment_chain:2"
misc_feature /note="assembly_fragment:00452"
fragment_chain:2"
misc_feature /note="assembly_fragment:00041"
fragment_chain:2"
BASE COUNT 68487 a 53697 c 52331 g 66069 t 1001 others
ORIGIN
Query Match 92.0% Score 18.4; DB 1; Length 241585;
Best Local Similarity 95.0%; Pred. No. 20;
Matches 19; Conservative 0; Mismatches 1; Gaps 0;
QY 1 TTAGTTGTAGTTAGTTAGTT 20
Db 164303 TTAGTTAGTTAGTTAGTT 164322
```

Search completed: April 6, 2003, 11:56:21  
Job time : 524.5 secs



```
source 1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="oligonucleotide"
BASE COUNT 4 a 0 c 4 g 12 t
ORIGIN

Query Match 100.0%; Score 20; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTAGTGTAGTTATTTAGTT 20
Db 1 TTAGTGTAGTTATTTAGTT 20

RESULT 2
AX057367 20 bp DNA linear PAT 17-JAN-2001
LOCUS
DEFINITION Sequence 1 from Patent WO0075304.
ACCESSION AX057367
VERSION AX057367.1 GI:12310108
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy,M., Sodoyer,R. and Tranoy,E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 1 14-DEC-2000;
Aventis Pasteur (FR)
FEATURES
source 1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="oligonucleotide"
BASE COUNT 5 a 0 c 3 g 12 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTAGTGTAGTTATTTAGTT 20
Db 1 TTAGTGTAGTTATTTAGTT 20

RESULT 3
AX057371 20 bp DNA linear PAT 17-JAN-2001
LOCUS
DEFINITION Sequence 5 from Patent WO0075304.
ACCESSION AX057371
VERSION AX057371.1 GI:12310112
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy,M., Sodoyer,R. and Tranoy,E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 5 14-DEC-2000;
Aventis Pasteur (FR)
FEATURES
source 1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="oligonucleotide"
BASE COUNT 4 a 0 c 3 g 13 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTAGTGTAGTTATTTAGTT 20
Db 1 TTAGTGTAGTTATTTAGTT 20
```

```
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTAGTGTAGTTATTTAGTT 20
Db 1 TTAGTGTAGTTATTTAGTT 20

RESULT 4
AX057375 20 bp DNA linear PAT 17-JAN-2001
LOCUS
DEFINITION Sequence 9 from Patent WO0075304.
ACCESSION AX057375
VERSION AX057375.1 GI:12310116
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy,M., Sodoyer,R. and Tranoy,E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 9 14-DEC-2000;
Aventis Pasteur (FR)
FEATURES
source 1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="oligonucleotide"
BASE COUNT 4 a 1 c 3 g 12 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTAGTGTAGTTATTTAGTT 20
Db 1 TTAGTGTAGTTATTTAGTT 20

RESULT 5
AX057380 20 bp DNA linear PAT 17-JAN-2001
LOCUS
DEFINITION Sequence 14 from Patent WO0075304.
ACCESSION AX057380
VERSION AX057380.1 GI:12310121
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy,M., Sodoyer,R. and Tranoy,E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 14 14-DEC-2000;
Aventis Pasteur (FR)
FEATURES
source 1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="oligonucleotide"
BASE COUNT 3 a 0 c 4 g 13 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTAGTGTAGTTATTTAGTT 20
Db 1 TTAGTGTAGTTATTTAGTT 20

RESULT 6
AX057381
```

LOCUS AX057381 20 bp DNA linear PAT 17-JAN-2001  
DEFINITION Sequence 15 from Patent W00075304.  
ACCESSION AX057381  
VERSION AX057381.1 GI:12310122  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Bachy,M., Sodoyer,R. and Trannoy,E.  
TITLE Immunostimulant oligonucleotide  
JOURNAL Patent: WO 0075304-A 15 14-DEC-2000;  
Aventis Pasteur (FR)  
FEATURES  
SOURCE 1..20  
location/Qualifiers  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="Oligonucleotide"  
BASE COUNT 3 a 1 c 4 g 12 t  
ORIGIN  
Query Match 92.0%; Score 18.4; DB 5; Length 20;  
Best Local Similarity 95.0%; Pred. No. 1.4e+03;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TTAGTTGTTAGTTATTAGTT 20  
|||||  
Db 1 TTAGTTGTTAGTTCTTACTT 20  
RESULT 7  
AX057382 20 bp DNA linear PAT 17-JAN-2001  
LOCUS AX057382  
DEFINITION Sequence 16 from Patent W00075304.  
ACCESSION AX057382  
VERSION AX057382.1 GI:12310123  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Bachy,M., Sodoyer,R. and Trannoy,E.  
TITLE Immunostimulant oligonucleotide  
JOURNAL Patent: WO 0075304-A 16 14-DEC-2000;  
Aventis Pasteur (FR)  
FEATURES  
SOURCE 1..20  
location/Qualifiers  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="Oligonucleotide"  
BASE COUNT 3 a 0 c 5 g 12 t  
ORIGIN  
Query Match 92.0%; Score 18.4; DB 5; Length 20;  
Best Local Similarity 95.0%; Pred. No. 1.4e+03;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TTAGTTGTTAGTTATTAGTT 20  
|||||  
Db 1 TTAGTTGTTAGTTGTTAGTT 20  
RESULT 8  
AF108143 2658 bp DNA linear BCT 17-FEB-2000  
LOCUS AF108143  
DEFINITION Anabaena PCC7120 Nt-Fe uptake hydrogenase small and large subunit genes, upstream sequence and unknown gene.  
ACCESSION AF108143  
VERSION AF108143.1 GI:5353551  
KEYWORDS  
SOURCE Nostoc sp. PCC 7120.  
ORGANISM Nostoc sp. PCC 7120  
Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
REFERENCE 1 (bases 1 to 2658)  
AUTHORS Golden,J.W.  
TITLE Anabaena sp. genomic sequence upstream of [nife] uptake hydrogenase small subunit (nups) gene and large subunit (nupL) gene  
JOURNAL unpublished  
REFERENCE 2 (bases 1 to 2658)  
AUTHORS Golden,J.W.  
TITLE Direct Submission  
JOURNAL Submitted (23-NOV-1998) Department of Biology, Texas A&M University, College Station, TX 77843-3258, USA  
FEATURES  
SOURCE 1..2658  
location/Qualifiers  
/organism="Nostoc sp., PCC 7120"  
/db\_xref="taxon:103690"  
/clone="pAM1268"  
/note="region upstream of Nt-Fe uptake hydrogenase small and large subunit genes" complement(510..1499)  
/codon\_start=1  
/transl\_table=1  
/product="unknown"  
/protein\_id="AAD42178.1"  
/db\_xref="GI:5353552"  
/translation="MTPSITDSTVKAALAAISSEATTEERKIQMLIELAQIQKPKPT PGDLMAVELQQAIRKMGDYSIMKRSQMGALSKSIDVGAELLQAKGYEEA LPVRLAAVEVAEAGMNFGLVLOSILPEFLARITDSIAAYOGAMGYVSODYPOEY AILANNVAIVLISMSGGSEONILFAGLAVOCTPEYATNTANKNALFNLDPDEKPE LQYLPSSHPENNIRATATATKATKVTCCGTPEYATNTANKNALFNLDPDEKPE LGNSQNLQARAYVOEAMEIFSQCOEMQAEVAQALHEVAALIRISMSNGK"  
BASE COUNT 838 a 664 c 423 g 733 t  
ORIGIN  
Query Match 92.0%; Score 18.4; DB 1; Length 2658;  
Best Local Similarity 95.0%; Pred. No. 2.9e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TTAGTTGTTAGTTATTAGTT 20  
|||||  
Db 2658 TTAGTTGTTAGTTGTTAGTT 2639  
RESULT 9  
ASU08013 2876 bp DNA linear BCT 30-NOV-1998  
LOCUS ASU08013  
DEFINITION Anabaena sp. [Nife] uptake hydrogenase small subunit (nups) and [Nife] uptake hydrogenase large subunit (nupL) genes, complete cds.  
ACCESSION U08013  
VERSION U08013.1 GI:3929403  
KEYWORDS  
SOURCE Anabaena sp.  
ORGANISM Anabaena sp.  
REFERENCE 1 (bases 1 to 2876)  
AUTHORS Carrasco,C.D., Buettner,J.A. and Golden,J.W.  
TITLE Programmed DNA rearrangement of a cyanobacterial nupL gene in heterocysts  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 92 (3), 791-795 (1995)  
MEDLINE 95146623  
PUBMED 7846053  
REFERENCE 2 (bases 1 to 2876)  
AUTHORS Carrasco,C.D., Holliday,S., Garcia,J.S. and Golden,J.W.  
TITLE xisc encodes a site-specific recombinase required for the developmentally regulated excision of the nupL element in Anabaena  
JOURNAL unpublished  
REFERENCE 3 (bases 1 to 2876)  
AUTHORS Golden,J.W.  
TITLE Direct Submission  
JOURNAL Submitted (24-MAR-1994) Biology, Texas A&M University, College Station, TX 77843-3258, USA  
REFERENCE 4 (bases 1 to 2876)  
AUTHORS Carrasco,C.D. and Golden,J.W.  
TITLE Direct Submission  
JOURNAL Submitted (30-NOV-1998) Biology, Texas A&M University, College

Station, TX 77843-3258, USA  
Sequence updated by submitter  
On Nov 30, 1998 this sequence version replaced gi:520458.  
COMMENT  
FEATURES  
source  
/organism="Anabaena sp."  
/strain="PCC 7120"  
/db\_xref="taxon:1167"  
/map="On BlnI-E and PstI-A between 0.4 and 0.9 Mb"  
/clone="PAM1268, PAM1314"  
/cell\_type="heterocyst"  
96.1072  
/gene="hups"  
96.1058  
/gene="hups"  
/codon\_start=1  
/transl\_table=11  
/product="[NifE] uptake hydrogenase small subunit"  
/protein\_id="AAC79877.1"  
/db\_xref="GI:3929404"  
/translation="MTNVLLOGAGSGNTMSFLNAEPTVCDLIADFGIKVLMHPSL  
GLEIDNLTLLMDCILGKIPDLIVFEGVYVNAPEGTGEMNFEADRPKMDLTDLAO  
AAEFVAVGDCATWGGIPAMPEPMSSEGLQFLKRGGGFLGODEPRTKSGDLINIG  
CAHPDWITQIIVATGRIADLDELNRPOFTFPTGTCTRVVHAKYATTAEP  
GOKGCLFTDLGCGPMTSSCNRIILMNVSSTKRAMPCLGCTEPEPFPLKPGTV  
FKQTIGVPEKELPGVSNKNYAVLTMVAKDAPKVAEDEFIV"  
1127.2722  
/gene="hupL"  
1127.2722  
/gene="hupL"  
/note="rearranged hupL gene on heterocyst chromosome"  
/codon\_start=1  
/transl\_table=11  
/product="[NifE] uptake hydrogenase large subunit"  
/protein\_id="AAC79878.1"  
/db\_xref="GI:520460"  
/translation="MTIKTLDISPVGRVSGDLVREIEDGRVNAWTHAELEFGEFI  
ILNGCPQAGLIVTPRICIGASHLTASMAIDTAMNTVPNNALIANLQIVETI  
OSIPRYEGLFALDITNKYRSRFEYDEAVRFSAYTGSELYTISKPEYELAF  
GGOWPSSVYPCGVACAPLIDITRAMALIEYFTNMLEPVMGLCSLREYETQYD  
DEMDLEADIKHRESDLGFWMGGLDIGDRGAGVGYKVSQVYLPHBKKYOKPTIEG  
RNAAMTKSGVDSFENRHTTMDHTTARENTTHAWDEGNADYHPDRTTKPHKATK  
DRNATSMSTAVLHODGRLEVGRLVAPLAKOLVAGGQHCESWQHDGFTLDLQKMGASI  
HLRQLARVHEIVKLQARCLREFVLNDPWYIKPEKGRGATGATASGSLCHMID  
IEGGKIKNOYVIAATTMVNGPRDSEGVGRPIEBALIGTIEDSRDEVEGVHARSFDS  
CIVCTVHAHDARTGAEELARFRA"  
1622.1637  
/gene="hupL"  
/note="hupL element recombination site on heterocyst  
chromosome"  
BASE COUNT 777 a 728 c 670 g 701 t  
ORIGIN  
Query Match 92.0%; Score 18.4; DB 1; Length 2876;  
Best Local Similarity 95.0%; Pred. No. 2.8e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 TTAGTCTAGTATATAGT 20  
||||| 1111111111111111  
Db 39 TTAGTCTAGTCTAGT 20

RESULT 10  
AF038558/c  
LOCUS 5568 bp DNA linear BCT 08-Apr-1998  
DEFINITION Fischerella PCC7605 photosystem I psaA (psaA) and photosystem I  
psaB (psaB) genes, complete cds.  
ACCESSION AF038558  
VERSION AF038558.1 GI:2827199  
KEYWORDS  
SOURCE Fischerella sp. PCC 7605.  
ORGANISM Fischerella sp. PCC 7605.  
Bacteria; Cyanobacteria; Stigonematales; Fischerella.

REFERENCE 1 (bases 1 to 5568)  
AUTHORS Sun,J., He,Z.Y., Nechushtai,R. and Chitnis,P.R.  
TITLE Molecular cloning of the psaA and psaB genes for the core proteins  
of photosystem I from the thermophilic cyanobacterium Mastigocladus  
lamninus (Accession No. AF038558) (PCR98-041)  
JOURNAL plant Physiol. 116 (3), 1192 (1998)  
REFERENCE 2 (bases 1 to 5568)  
AUTHORS Sun,J., He,Z.Y., Nechushtai,R. and Chitnis,P.R.  
TITLE Direct Submissions  
JOURNAL Submitted (13-DEC-1997) Biochemistry and Biophysics, Iowa State  
University, 4156 Mol. Biol. Bldg., Ames, IA 50011, USA  
FEATURES  
source  
/organism="Fischerella sp. PCC 7605"  
/strain="PCC7605"  
/db\_xref="taxon:60455"  
/note="submitted as Mastigocladus lamninus"  
593..2851  
/gene="psaA"  
593..2851  
/gene="psaA"  
/note="binds chlorophylls, iron-sulfur centers and  
phyllonones"  
/codon\_start=1  
/transl\_table=11  
/product="photosystem I psaA"  
/protein\_id="AAC12866.1"  
/db\_xref="GI:2827200"  
/translation="MTISPPREKKARIVYNDPVPYSFEIMAKRGHPRTLARPK  
TTTWIMNLALAHDPRTSHSDELSRKFPAHFGHLAVVTMLSMITRGRFSYE  
AMISDLVHVPASQVWPVIGODTLNGDVGDFGFGIQTSGLOIRMGGLTNSFLY  
VAIGGLVLAGLEFAGWPHYKRAKLEPMESBMLNHLAVLLGCGSLMAGHLI  
HVSAPTNKLADGVSVDIPLPHFELTLELITVIGHWYRTMMGSGHSIKELNNHGPFL  
TFKGLNPTVGGMLTDLSDHNLDAIVLEITIGHWYRTMMGSGHSIKELNNHGPFL  
GGHKGLEYENMTSMHAOLATNLAFGLSTITVAHWYAMPYRATYATQOLCET  
HHMIGGFLVGAANHTFTPRDIPVYNNQNVADVYRHRDITSHINWCYIEFG  
HSFGVLVNDTRALGRPODMFSDAIDQVFAQVQVHITLAPSTAPNLEPAST  
AGCGIVAVAGGVVAMPVIALGADPLVHNIHAFTHVTLILKGLFARNRLLPDK  
ANLGEFFPCDGRGTCGVSMDFVLEGLFPMYNSISIVIFHSKMQSDVMGTVD  
ACNVHTITGNNAAOSALTINGMLRDLMAQSVQVINSYGSALSAYGIMLGHIWAF  
SIMPLESGRGVYVQELIESIVMAHNKIKVAPAIQPRALSTIQGRVAVHYLLGIATY  
NAFPAHILSLIC"  
2971..3202  
/gene="psaB"  
2971..3202  
/gene="psaB"  
/note="binds chlorophylls, iron-sulfur centers and  
phyllonones"  
/codon\_start=1  
/transl\_table=11  
/product="photosystem I psaB"  
/protein\_id="AAC12867.1"  
/db\_xref="GI:2827201"  
/translation="MATKPKFSODLAODPTTRIMYAMATGNDPESHGDMENLYO  
KIFATRFVAVILFTMASSLLFRVANQFEDQKIDPHIRPRAHNDPOGRKRAIE  
ATQOGGANPVINISGLYHMMYTGMRNNNDLQKAVFLILASVFLFAGLHLQPK  
FRPSLSWFKSAEPLRNHLGLFVSSLSMAQCHLHVAIPESRSGVGNNTLSTPH  
PAGIDPFEGNAGVYASDDPTANVFGTQGGAGTALITFLGFGFHQTESLMTDANH  
HLAIVYFTIAGHMYRTNFGHSIKEMNSKAGVPCVNSGQFNLPHQGYDYNN  
SLHFGILHILAIGTTSIVAOHWKAMPYATLADYTOALATYTHOVALTELMGA  
FAHGALEWVDYDPEONKGNVLDRLYKHEALISHSVSLFLGRTTLDLYHNDVY  
AGTPEKQILIEPVFAQFLQASHGKVLGLVNLSPDSVATATAPNGVNWQGLD  
AINGNSLFTLIGPDLVHHAFAIATHTVLLVKGALDARSGLMDPKDFGAF  
PCDGRGTCGCOISADAFYLATFNALNVGTVFVHMKHLIGOGNVAQDNESSTY  
LKGWPRDYAMANSADLNGVYNNKNNISVMAWMLFGLHLVWATGMFLISRGYWOE  
LIETLVMAHERPLANLVWKKPVALSTIVQARVLGLHTFTYGYVLTLYAFLIASTAG  
KFC"  
BASE COUNT 1378 a 1396 c 1235 g 1559 t  
ORIGIN  
Query Match 92.0%; Score 18.4; DB 1; Length 5568;  
Best Local Similarity 95.0%; Pred. No. 2.2e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTTGTTAGTTATTAGTT 20  
|||||  
Db 2924 TTAGTTGTTAGTTATTAGTT 2905

RESULT 11  
AX251168 11029 bp DNA linear PAT 05-OCT-2001  
LOCUS Sequence 136 from Patent WO0168912.  
ACCESSION AX251168  
VERSION AX251168.1 GI:15984591  
KEYWORDS  
ORGANISM synthetic construct.  
SOURCE synthetic construct.  
REFERENCE 1 (bases 1 to 11029)  
AUTHORS Olek,A., Piepenbrock,C. and Berlin,K.  
TITLE Diagnosis of diseases associated with tumor suppressor genes and oncogenes  
JOURNAL Patent: WO 0168912-A 136 20-SEP-2001;  
EpiGenomics AG (DE)  
FEATURES  
source Location/Qualifiers  
1..11029  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="chemically treated genomic DNA (Homo sapiens)"

BASE COUNT 2865 a 205 c 2388 g 5571 t  
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 11029;  
Best Local Similarity 95.0%; Pred. No. 1.6e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTTGTTAGTTATTAGTT 20  
|||||  
Db 7283 TTAGTTGTTAGTTATTAGTT 7302

RESULT 12  
AX348425 16258 bp DNA linear PAT 06-FEB-2002  
LOCUS Sequence 120 from Patent WO0202806.  
DEFINITION AX348425  
ACCESSION AX348425  
VERSION AX348425.1 GI:18614461  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1  
AUTHORS Olek,A., Piepenbrock,C. and Berlin,K.  
TITLE Method and nucleic acids for pharmacogenomic methylation analysis  
JOURNAL Patent: WO 0202806-A 120 10-JAN-2002;  
EpiGenomics AG (DE)  
FEATURES  
source Location/Qualifiers  
1..16258  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="chemically treated genomic DNA (Homo sapiens)"

BASE COUNT 4242 a 330 c 3801 g 7885 t  
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 16258;  
Best Local Similarity 95.0%; Pred. No. 1.6e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTTGTTAGTTATTAGTT 20  
|||||  
Db 16029 TTAGTTGTTAGTTATTAGTT 16048

RESULT 13  
AX348808

LOCUS AX348808 16258 bp DNA linear PAT 06-FEB-2002  
DEFINITION Sequence 266 from Patent WO0202807.  
ACCESSION AX348808  
VERSION AX348808.1 GI:18614843  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1  
AUTHORS Olek,A., Piepenbrock,C. and Berlin,K.  
TITLE Diagnosis of diseases associated with cell signalling  
JOURNAL Patent: WO 0202807-A 266 10-JAN-2002;  
EpiGenomics AG (DE)  
FEATURES  
source Location/Qualifiers  
1..16258  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="chemically treated genomic DNA (Homo sapiens)"

BASE COUNT 4242 a 330 c 3801 g 7885 t  
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 16258;  
Best Local Similarity 95.0%; Pred. No. 1.6e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTTGTTAGTTATTAGTT 20  
|||||  
Db 16029 TTAGTTGTTAGTTATTAGTT 16048

RESULT 14  
AP004902/c 123078 bp DNA linear PAT 19-JUL-2002  
LOCUS Lotus japonicus genomic DNA, chromosome 2, clone: LjT04G24, TM0060, complete sequence.  
DEFINITION AP004902  
ACCESSION AP004902  
VERSION AP004902.1 GI:21907918  
KEYWORDS  
SOURCE HTG.  
ORGANISM Lotus japonicus  
Lotus japonicus DNA, clone: LjT library clone: LjT04G24.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.

REFERENCE 1  
AUTHORS Kaneko,T., Nakamura,Y., Asamizu,E., Kato,T., Sato,S. and Tabata,S.  
TITLE Structural Analysis of a Lotus japonicus Genome. I. Sequence Features and Mapping of Sixty-six TAC clones which cover the 6.7 Mb Regions of the Genome  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 123078)  
AUTHORS Nakamura,Y.  
TITLE Direct Submission  
JOURNAL Submitted (26-MAR-2002) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research, 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: ynk@kazusa.or.jp, URL: http://www.kazusa.or.jp, Tel: 81-438-52-3935, Fax: 81-438-52-3934)

FEATURES  
source Location/Qualifiers  
1..123078  
/organism="Lotus japonicus"  
/db\_xref="taxon:34305"  
/chromosome="2"  
/clone="LjT04G24"  
/clone="LjT library"  
/note="TAC clone: TM0060"

BASE COUNT 40721 a 22328 c 21292 g 38737 t  
ORIGIN

Query Match 92.0%; Score 18.4; DB 7; Length 123078;  
Best Local Similarity 95.0%; Pred. No. 81;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTGTAGTATTAGTT 20  
 |||||  
 Db 40067 TTAGTGTAGTATTAGTT 40048

RESULT 15  
 AC010232/c 165920 bp DNA linear PRI 26-FEB-2002  
 LOCUS Homo sapiens chromosome 5 clone CTC-316C21, complete sequence.  
 DEFINITION AC010232  
 ACCESSION AC010232  
 VERSION AC010232.10 GI:18921277  
 KEYWORDS HTG.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 165920)  
 DOE Joint Genome Institute and Stanford Human Genome Center.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 165920)  
 DOE Joint Genome Institute.  
 Direct Submission  
 Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint  
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 3 (bases 1 to 165920)  
 DOE Joint Genome Institute and Stanford Human Genome Center.  
 Direct Submission  
 Submitted (26-FEB-2002) DOE Joint Genome Institute, 2800 Mitchell  
 Drive, Walnut Creek, CA 94598, USA  
 On Feb 26, 2002 this sequence version replaced gi:15290321.  
 Draft Sequence Produced by DOE Joint Genome Institute  
 www.jgi.doe.gov  
 Finishing Completed at Stanford Human Genome Center  
 www.shgc.stanford.edu  
 Quality: Phrap Quality >=40 99.9% of Sequence:  
 Estimated Total Number of Errors is 0.1.  
 Location/Qualifiers  
 1. 165920  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="5"  
 /clone="CTC-316C21"

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

FEATURES  
 source

BASE COUNT 45866 a 32867 c 34867 g 52500 t  
 ORIGIN

Query Match 92.0%; Score 18.4; DB 8; Length 165920;  
 Best Local Similarity 95.0%; Pred. No. 74;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTGTAGTATTAGTT 20  
 |||||  
 Db 50747 TTAGTGTAGTATTAGTT 50728

Search completed: April 6, 2003, 11:57:43  
 Job time : 305 secs



GenCore version 5.1.4-P5-4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 11:21:32 ; Search time 151 Seconds  
(without alignments)  
298.278 Million cell updates/sec

Title: US-09-980-265-12

Perfect score: 20  
Sequence: 1 ttagttagttagttagt 20

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

N.Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	AAF28884	Immunostimulatory
2	18.4	92.0	20	AAF28872	Immunostimulatory
3	18.4	92.0	20	AAF28876	Immunostimulatory
4	18.4	92.0	20	AAF28880	Immunostimulatory
5	18.4	92.0	20	AAF28885	Immunostimulatory
6	18.4	92.0	20	AAF28886	Immunostimulatory
7	18.4	92.0	20	AAF28887	Immunostimulatory
8	18.4	92.0	749	ABO21728	Oligonucleotide fo
9	18.4	92.0	749	ABO21729	Oligonucleotide fo

10	18.4	92.0	11029	22	AA546414	Tumour suppressor
11	18.4	92.0	16258	24	ABL70376	Chemically treated
12	18.4	92.0	16258	24	ABK40038	Human chemically p
13	18	90.0	27435	22	AAK56476	Human immune/haema
14	17.4	87.0	20	22	AAE28871	Human immune/haema
15	17.4	87.0	2526	23	ABL13576	Genetic immunostim
16	17.4	87.0	4037	23	ABL11048	Drosophila melanog
17	17.4	87.0	5293	24	ABL13658	Human immune syste
18	17.4	87.0	7238	24	ABL32448	Human immune syste
19	17.4	87.0	9507	24	ABN80243	Human immune syste
20	17.4	87.0	9728	24	ABK40092	Human chemically m
21	17.4	87.0	9728	24	ABL33903	Human immune syste
22	17.4	87.0	15698	24	ABL34141	Human immune syste
23	17.4	87.0	15861	24	ABL33255	Human immune syste
24	17.4	87.0	6419	24	ABL33358	Human immune syste
25	17	85.0	6419	24	ABK28377	Human immune syste
26	17	85.0	6522	24	ABL33050	DNA transcription
27	17	85.0	11050	24	ABL43386	Human immune syste
28	17	85.0	12986	22	AA546553	Human polynucleoti
29	16.8	84.0	20	22	AAF28873	Tumour suppressor
30	16.8	84.0	20	22	AAF28874	Immunostimulatory
31	16.8	84.0	20	22	AAF28875	Immunostimulatory
32	16.8	84.0	20	22	AAF28877	Immunostimulatory
33	16.8	84.0	20	22	AAF28878	Immunostimulatory
34	16.8	84.0	20	22	AAF28879	Immunostimulatory
35	16.8	84.0	20	22	AAF28881	Immunostimulatory
36	16.8	84.0	20	22	AAF28882	Immunostimulatory
37	16.8	84.0	20	22	AAF28883	Immunostimulatory
38	16.8	84.0	511	24	ABO20686	Immunostimulatory
39	16.8	84.0	511	24	ABO20687	Oligonucleotide fo
40	16.8	84.0	529	24	ABO21336	Oligonucleotide fo
41	16.8	84.0	529	24	ABO21337	Oligonucleotide fo
42	16.8	84.0	548	24	ABN64773	Oligonucleotide fo
43	16.8	84.0	659	24	ABO49920	Human cancer relat
44	16.8	84.0	659	24	ABO49921	Oligonucleotide fo
45	16.8	84.0	717	24	ABO24820	Oligonucleotide fo

#### ALIGNMENTS

RESULT 1	AAF28884	standard: DNA; 20 Bp.
ID	AAF28884:	
AC	09-MAY-2001 (first entry)	
XX	Immunostimulatory oligonucleotide #13 as vaccine adjuvant.	
DT		
DE		
XX	Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;	
KW	prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;	
KW	cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;	
KW	phosphorothioate; ss.	
XX		
OS	Synthetic.	
XX		
FH	Key	Location/Qualifiers
FT	modified_base	1..20
FT		/*tag= a
FT		/note= "contain phosphorothioate internucleotide linkages"
FT		
XX		
PN	WO200075304-A1.	
XX		
PD	14-DEC-2000.	
XX		
PF	08-JUN-2000; 2000WO-FR01566.	
XX		
PR	08-JUN-1999; 99FR-0007457.	
PR	06-AUG-1999; 99FR-0010378.	
XX		

PA (AVER ) AVENTIS PASTEUR.  
XX  
PI Bachy M, Sodayer R, Trannoy E;  
XX  
DR WPI; 2001-041317/05.  
XX  
PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in  
PT vaccines for human use, induce lymphocyte proliferation and cytokine  
PT secretion -  
XX  
PS Example 4; Page 13; 30pp; French.  
XX  
CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an  
CC immunostimulatory oligonucleotide of the invention which contains at  
CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The  
CC oligonucleotides do not contain any CG dinucleotides in which C are  
CC unethylated. The oligonucleotides are used as human immunostimulants  
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.  
CC They induce proliferation of human lymphocytes, induce secretion of  
CC cytokines, especially Interleukin-10 or Interferon-gamma and increase  
CC expression of the CD86 activation marker or the CD25 cytokine receptor  
CC on human B lymphocytes.  
CC  
XX  
SQ Sequence 20 BP; 4 A; 0 C; 4 G; 12 T; 0 other;  
Query Match 100.0%; Score 20; DB 22; Length 20;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 TTAGTGTAGTTATTAGTT 20  
1 TTAGTGTAGTTATTAGTT 20  
DB 1 TTAGTGTAGTTATTAGTT 20  
RESULT 2  
AAF28872  
ID AAF28872 standard; DNA: 20 BP.  
XX  
AC AAF28872;  
XX  
DT 09-MAY-2001 (first entry)  
XX  
DE Immunostimulatory oligonucleotide #1 as vaccine adjuvant.  
XX  
KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;  
KW Propylactic vaccine; human lymphocyte proliferation; interferon-gamma;  
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;  
KW phosphorothioate; ss.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT modified\_base 1..20  
FT /\*tag= a  
FT /note= "contain phosphorothioate internucleotide  
FT linkages"  
XX  
PN WO200075304-A1.  
XX  
PD 14-DEC-2000.  
XX  
PF 08-JUN-2000; 2000MO-FR01566.  
XX  
PR 08-JUN-1999; 99FR-0007457.  
PR 06-AUG-1999; 99FR-0010378.  
XX  
XX (AVER ) AVENTIS PASTEUR.  
XX  
PI Bachy M, Sodayer R, Trannoy E;  
XX  
DR WPI; 2001-041317/05.  
XX  
PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in

PT vaccines for human use, induce lymphocyte proliferation and cytokine  
PT secretion -  
XX  
PS Example 4; Page 13; 30pp; French.  
XX  
CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an  
CC immunostimulatory oligonucleotide of the invention which contains at  
CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The  
CC oligonucleotides do not contain any CG dinucleotides in which C are  
CC unethylated. The oligonucleotides are used as human immunostimulants  
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.  
CC They induce proliferation of human lymphocytes, induce secretion of  
CC cytokines, especially Interleukin-10 or Interferon-gamma and increase  
CC expression of the CD86 activation marker or the CD25 cytokine receptor  
CC on human B lymphocytes.  
CC  
XX  
SQ Sequence 20 BP; 5 A; 0 C; 3 G; 12 T; 0 other;  
Query Match 92.0%; Score 18.4; DB 22; Length 20;  
Best Local Similarity 95.0%; Pred. No. 97;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 TTAGTGTAGTTATTAGTT 20  
1 TTAGTGTAGTTATTAGTT 20  
DB 1 TTAGTGTAGTTATTAGTT 20  
RESULT 3  
AAF28876  
ID AAF28876 standard; DNA: 20 BP.  
XX  
AC AAF28876;  
XX  
DT 09-MAY-2001 (first entry)  
XX  
DE Immunostimulatory oligonucleotide #5 as vaccine adjuvant.  
XX  
KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;  
KW Propylactic vaccine; human lymphocyte proliferation; interferon-gamma;  
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;  
KW phosphorothioate; ss.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT modified\_base 1..20  
FT /\*tag= a  
FT /note= "contain phosphorothioate internucleotide  
FT linkages"  
XX  
PN WO200075304-A1.  
XX  
PD 14-DEC-2000.  
XX  
PF 08-JUN-2000; 2000MO-FR01566.  
XX  
PR 08-JUN-1999; 99FR-0007457.  
PR 06-AUG-1999; 99FR-0010378.  
XX  
XX (AVER ) AVENTIS PASTEUR.  
XX  
PI Bachy M, Sodayer R, Trannoy E;  
XX  
DR WPI; 2001-041317/05.  
XX  
PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in  
PT vaccines for human use, induce lymphocyte proliferation and cytokine  
PT secretion -  
XX  
PS Example 4; Page 13; 30pp; French.  
XX  
CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an  
CC immunostimulatory oligonucleotide of the invention which contains at

CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The  
 CC oligonucleotides do not contain any CG dinucleotides in which C are  
 CC unethyiated. The oligonucleotides are used as human immunostimulants  
 CC and as adjuvants in therapeutic and prophylactic vaccines for human use.  
 CC They induce proliferation of human lymphocytes, induce secretion of  
 CC cytokines, especially interleukin-10 or interferon-gamma and increase  
 CC expression of the CD86 activation marker or the CD25 cytokine receptor  
 CC on human B lymphocytes.

XX Sequence 20 BP: 4 A; 0 C; 3 G; 13 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 20;

Best Local Similarity 95.0%; Pred. No. 97;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTGTAGTTATTAGTT 20  
 Db 1 TTAGTGTAGTTATTAGTT 20

#### RESULT 4

ID AAF28880 standard; DNA; 20 BP.

XX AAF28880;

DT 09-MAY-2001 (first entry)

XX Immunostimulatory oligonucleotide #9 as vaccine adjuvant.

XX Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;

KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;

KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;

XX phosphorothioate; ss.

XX Synthetic.

FT Key Location/Qualifiers  
 FT modified\_base 1..20  
 FT /tag= a  
 FT /note= "contain phosphorothioate internucleotide  
 FT linkages"

XX WO200075304-A1.

XX 14-DEC-2000.

XX 08-JUN-2000; 2000WO-FR01566.

XX 08-JUN-1999; 99FR-0007457.

XX 06-AUG-1999; 99FR-0010376.

XX (AVET ) AVENTIS PASTEUR.

XX Bachy M, Sodoyer R, Tranoy E;

XX WPI; 2001-041317/05.

PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in  
 PT vaccines for human use, induce lymphocyte proliferation and cytokine  
 PT secretion -  
 XX Example 4; Page 13; 30pp; French.

CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an  
 CC immunostimulatory oligonucleotide of the invention which contains at  
 CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The  
 CC oligonucleotides do not contain any CG dinucleotides in which C are  
 CC unethyiated. The oligonucleotides are used as human immunostimulants  
 CC and as adjuvants in therapeutic and prophylactic vaccines for human use.  
 CC They induce proliferation of human lymphocytes, induce secretion of  
 CC cytokines, especially interleukin-10 or interferon-gamma and increase  
 CC expression of the CD86 activation marker or the CD25 cytokine receptor

CC on human B lymphocytes.

XX Sequence 20 BP: 4 A; 1 C; 3 G; 12 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 20;

Best Local Similarity 95.0%; Pred. No. 97;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTGTAGTTATTAGTT 20  
 Db 1 TTAGTGTAGTTATTAGTT 20

#### RESULT 5

ID AAF28885 standard; DNA; 20 BP.

XX AAF28885;

DT 09-MAY-2001 (first entry)

XX Immunostimulatory oligonucleotide #14 as vaccine adjuvant.

XX Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;

KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;

KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;

XX phosphorothioate; ss.

XX Synthetic.

FT Key Location/Qualifiers  
 FT modified\_base 1..20  
 FT /tag= a  
 FT /note= "contain phosphorothioate internucleotide  
 FT linkages"

XX WO200075304-A1.

XX 14-DEC-2000.

XX 08-JUN-2000; 2000WO-FR01566.

XX 08-JUN-1999; 99FR-0007457.

XX 06-AUG-1999; 99FR-0010376.

XX (AVET ) AVENTIS PASTEUR.

XX Bachy M, Sodoyer R, Tranoy E;

XX WPI; 2001-041317/05.

PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in  
 PT vaccines for human use, induce lymphocyte proliferation and cytokine  
 PT secretion -  
 XX Example 4; Page 13; 30pp; French.

CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an  
 CC immunostimulatory oligonucleotide of the invention which contains at  
 CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The  
 CC oligonucleotides do not contain any CG dinucleotides in which C are  
 CC unethyiated. The oligonucleotides are used as human immunostimulants  
 CC and as adjuvants in therapeutic and prophylactic vaccines for human use.  
 CC They induce proliferation of human lymphocytes, induce secretion of  
 CC cytokines, especially interleukin-10 or interferon-gamma and increase  
 CC expression of the CD86 activation marker or the CD25 cytokine receptor  
 CC on human B lymphocytes.

XX Sequence 20 BP: 3 A; 0 C; 4 G; 13 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 20;

Best Local Similarity 95.0%; Pred. No. 97;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTGTCTAGTTATTAGTT 20  
 |||||  
 DB 1 TTAGTGTCTAGTTATTAGTT 20

## RESULT 6

AAAF28886  
 ID AAF28886 standard; DNA; 20 BP.

AC AAF28886;

DT 09-MAY-2001 (first entry)

DE Immunostimulatory oligonucleotide #15 as vaccine adjuvant.

KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;  
 KW propylactic vaccine; human lymphocyte proliferation; interferon-gamma;  
 KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;  
 KW phosphorothioate; ss.

OS Synthetic.

Key Location/Qualifiers  
 modified\_base 1..20  
 FT /tag= a  
 FT /note= "contain phosphorothioate internucleotide linkages"

MO200075304-A1.

PD 14-DEC-2000.

PF 08-JUN-2000; 2000MO-FR01566.

PR 08-JUN-1999; 99FR-0007457.

PR 06-AUG-1999; 99FR-0010378.

PA (AVET ) AVENTIS PASTEUR.

PI Bachy M, Sodayer R, Tranroy E;

DR WPI; 2001-041317/05.

PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in  
 vaccines for human use, induce lymphocyte proliferation and cytokine  
 secretion -

XX Example 4; Page 13; 30pp; French.

CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an  
 CC immunostimulatory oligonucleotide of the invention which contains at  
 CC least one sequence 5'-TTN1N2TT-3' where N1 and N2 are A, T, C or G. The  
 CC oligonucleotides do not contain any CG dinucleotides in which C are  
 CC unmethylated. The oligonucleotides are used as human immunostimulants  
 CC and as adjuvants in therapeutic and prophylactic vaccines for human use.  
 CC They induce proliferation of human lymphocytes, induce secretion of  
 CC cytokines, especially interleukin-10 or interferon-gamma and increase  
 CC expression of the CD86 activation marker or the CD25 cytokine receptor  
 CC on human B lymphocytes.

XX Sequence 20 BP; 3 A; 1 C; 4 G; 12 T; 0 other;

QY Query Match 92.0%; Score 18.4; DB 22; Length 20;

Best Local Similarity 95.0%; Pred. No. 97;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTGTCTAGTTATTAGTT 20  
 |||||  
 DB 1 TTAGTGTCTAGTTATTAGTT 20

## RESULT 7

AAAF28887  
 ID AAF28887 standard; DNA; 20 BP.

AC AAF28887;

DT 09-MAY-2001 (first entry)

DE Immunostimulatory oligonucleotide #16 as vaccine adjuvant.

KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;  
 KW propylactic vaccine; human lymphocyte proliferation; interferon-gamma;  
 KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;  
 KW phosphorothioate; ss.

OS Synthetic.

Key Location/Qualifiers  
 modified\_base 1..20  
 FT /tag= a  
 FT /note= "contain phosphorothioate internucleotide linkages"

MO200075304-A1.

PD 14-DEC-2000.

PF 08-JUN-2000; 2000MO-FR01566.

PR 08-JUN-1999; 99FR-0007457.

PR 06-AUG-1999; 99FR-0010378.

PA (AVET ) AVENTIS PASTEUR.

PI Bachy M, Sodayer R, Tranroy E;

DR WPI; 2001-041317/05.

PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in  
 vaccines for human use, induce lymphocyte proliferation and cytokine  
 secretion -

XX Example 4; Page 13; 30pp; French.

CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an  
 CC immunostimulatory oligonucleotide of the invention which contains at  
 CC least one sequence 5'-TTN1N2TT-3' where N1 and N2 are A, T, C or G. The  
 CC oligonucleotides do not contain any CG dinucleotides in which C are  
 CC unmethylated. The oligonucleotides are used as human immunostimulants  
 CC and as adjuvants in therapeutic and prophylactic vaccines for human use.  
 CC They induce proliferation of human lymphocytes, induce secretion of  
 CC cytokines, especially interleukin-10 or interferon-gamma and increase  
 CC expression of the CD86 activation marker or the CD25 cytokine receptor  
 CC on human B lymphocytes.

XX Sequence 20 BP; 3 A; 0 C; 5 G; 12 T; 0 other;

QY Query Match 92.0%; Score 18.4; DB 22; Length 20;

Best Local Similarity 95.0%; Pred. No. 97;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTGTCTAGTTATTAGTT 20  
 |||||  
 DB 1 TTAGTGTCTAGTTATTAGTT 20

## RESULT 8

ABO21728  
 ID ABO21728 standard; DNA; 749 BP.

AC ABO21728;

DT 12-JUL-2002 (first entry)

DE oligonucleotide for detecting cytosine methylation SEQ ID NO 8319.  
 XX Human: cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
 KW drug; side effect; cancer; central nervous system; cardiovascular;  
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;  
 KW SNP; cell differentiation; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200218632-A2.  
 XX  
 PD 07-MAR-2002.  
 XX  
 PF 01-SEP-2001; 2001WO-EPI0074.  
 XX  
 PR 01-SEP-2000; 2000DE-1043826.  
 XX  
 PR 05-SEP-2000; 2000DE-1044543.  
 XX  
 PA (EPIC-) EPIGENOMICS AG.  
 XX  
 PI Olek A, Piepenbrock C, Berlin K, Guetig D;  
 XX  
 DR WPI; 2002-371829/40.  
 XX  
 PT Determining the degree of cytosine methylation in genomic DNA, useful  
 PT for diagnosis and prognosis, comprises selective hybridization of  
 PT amplicons from chemically treated DNA -  
 XX  
 PS Claim 12; 56pp + Sequence Listing; 56pp; German.  
 XX  
 CC This invention describes a novel method for determining the degree of  
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
 CC genomic sample of DNA. The sample is treated chemically to convert  
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
 CC DNA that contains the target C is amplified to form a labeled amplicon.  
 CC The amplicon is hybridised to two classes, each with at least one  
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers  
 CC and the degree of hybridisation to both classes is determined from the  
 CC label on the amplicon. From the ratio of labels hybridised to the two  
 CC classes of oligomers, the degree of methylation is calculated. The method  
 CC is used: (i) for diagnosis and/or prognosis of side effects of  
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders  
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory  
 CC systems etc., particularly by detecting mutations or single nucleotide  
 CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue  
 CC types and for investigating cell differentiation. The method allows the  
 CC methylation status of many C residues to be determined simultaneously.  
 CC AB013410-AB054121 represent genomic DNA sequences used to illustrate the  
 CC method for determining the degree of cytosine methylation described in  
 CC the disclosure of the invention.  
 CC  
 XX  
 SQ Sequence 749 BP; 173 A; 54 C; 199 G; 322 T; 1 other;  
 XX  
 QY Query Match 92.0%; Score 18.4; DB 24; Length 749;  
 QY Best Local Similarity 95.0%; Pred. NO. 80;  
 QY Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 TTAGTGTGTTAGTTATTAGTT 20  
 QY ||||| ||||| ||||| |||||  
 DB 347 TTAGTTTATTAGTTATTAGTT 366  
 XX  
 RESULT 9  
 ABQ21729/c  
 ID ABQ21729 standard; DNA; 749 BP.  
 XX  
 AC ABQ21729;  
 XX  
 DT 12-JUL-2002 (first entry)  
 XX  
 DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 8320.  
 XX  
 KW Human: cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;

KW drug; side effect; cancer; central nervous system; cardiovascular;  
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;  
 KW SNP; cell differentiation; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200218632-A2.  
 XX  
 PD 07-MAR-2002.  
 XX  
 PF 01-SEP-2001; 2001WO-EPI0074.  
 XX  
 PR 01-SEP-2000; 2000DE-1043826.  
 XX  
 PR 05-SEP-2000; 2000DE-1044543.  
 XX  
 PA (EPIC-) EPIGENOMICS AG.  
 XX  
 PI Olek A, Piepenbrock C, Berlin K, Guetig D;  
 XX  
 DR WPI; 2002-371829/40.  
 XX  
 PT Determining the degree of cytosine methylation in genomic DNA, useful  
 PT for diagnosis and prognosis, comprises selective hybridization of  
 PT amplicons from chemically treated DNA -  
 XX  
 PS Claim 12; 56pp + Sequence Listing; 56pp; German.  
 XX  
 CC This invention describes a novel method for determining the degree of  
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
 CC genomic sample of DNA. The sample is treated chemically to convert  
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
 CC DNA that contains the target C is amplified to form a labeled amplicon.  
 CC The amplicon is hybridised to two classes, each with at least one  
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers  
 CC and the degree of hybridisation to both classes is determined from the  
 CC label on the amplicon. From the ratio of labels hybridised to the two  
 CC classes of oligomers, the degree of methylation is calculated. The method  
 CC is used: (i) for diagnosis and/or prognosis of side effects of  
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders  
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory  
 CC systems etc., particularly by detecting mutations or single nucleotide  
 CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue  
 CC types and for investigating cell differentiation. The method allows the  
 CC methylation status of many C residues to be determined simultaneously.  
 CC AB013410-AB054121 represent genomic DNA sequences used to illustrate the  
 CC method for determining the degree of cytosine methylation described in  
 CC the disclosure of the invention.  
 CC  
 XX  
 SQ Sequence 749 BP; 322 A; 199 C; 54 G; 173 T; 1 other;  
 XX  
 QY Query Match 92.0%; Score 18.4; DB 24; Length 749;  
 QY Best Local Similarity 95.0%; Pred. NO. 80;  
 QY Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 TTAGTGTGTTAGTTATTAGTT 20  
 QY ||||| ||||| ||||| |||||  
 DB 403 TTAGTTTATTAGTTATTAGTT 384  
 XX  
 RESULT 10  
 AAS46414  
 ID AAS46414 standard; DNA; 11029 BP.  
 XX  
 AC AAS46414;  
 XX  
 DT 18-DEC-2001 (first entry)  
 XX  
 DE Tumour suppressor gene derived chemically modified sequence #136.  
 XX  
 KW Human: tumour suppressor gene; oncogene; antitumour; cytostatic;  
 KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;  
 KW cytosine methylation; ds.

```
OS Homo sapiens.
XX
XX WO200168912-A2.
XX
XX 20-SEP-2001.
XX
XX 15-MAR-2001: 2001WO-EP02955.
XX
XX 15-MAR-2000: 2000DE-1013847.
XX
XX 06-APR-2000: 2000DE-1019058.
XX
XX 07-APR-2000: 2000DE-1019173.
XX
XX 30-JUN-2000: 2000DE-1032529.
XX
XX 01-SEP-2000: 2000DE-1043826.
XX
XX (EPIC-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI: 2001-602752/68.
XX
XX Fragments of chemically modified genes associated with tumour suppressor
XX genes and oncogenes, useful in designing primers and probes for
XX analysing diseases associated with cytosine methylation state e.g.
XX cancer
XX
XX Claim 1: SEQ ID NO 136; 27pp; English.
XX
XX The invention relates to a nucleic acid comprising a sequence of 18
XX bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
XX bisulphite, of genes associated with tumour suppression and
XX oncogenes having a sequence taken from 536 (actually 533 since
XX numbers 408, 458 and 500 are missing from the sequence listing) sequences
XX (5s) and sequences complementary to (5s). The nucleic acid may be a
XX peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
XX form part of a set of probes for detecting the cytosine methylation state
XX and/or single nucleotide polymorphisms and also to be used in an
XX array for analysing diseases associated with CpG dinucleotides e.g.
XX cancers and tumours. The probes can also be used in a method for
XX ascertaining genetic and/or epigenetic parameters for the diagnosis
XX and/or therapy of existing diseases or the predisposition to specific
XX diseases, by analysing cytosine methylations. The parameters may be
XX compared to another set of genetic and/or epigenetic parameters, the
XX differences serving as basis for diagnosis and/or prognosis events which
XX are disadvantageous to patients. The present sequence is one of the
XX 533 genomic sequences derived from tumour suppressor genes and
XX oncogenes. Sequences with even numbered Seq ID numbers are the
XX complementary sequence of the corresponding odd numbered sequence (e.g.
XX ID 2 and ID1, ID 536 and ID 535, except for those whose partner sequence
XX is missing).
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic
XX format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 11029 BP; 2865 A; 205 C; 2388 G; 5571 T; 0 other;
XX
XX Query Match 92.0%; Score 18.4; DB 22; Length 11029;
XX Best Local Similarity 95.0%; Pred. No. 70;
XX Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 TTAGTGTAGTTATTAGTT 20
XX ||||||||||||||||
XX DB 7283 TTAGTGTAGTTATTATT 7302
XX
XX RESULT 11
XX ABL70376
XX ID ABL70376 standard; DNA: 16258 BP.
XX
XX ABL70376;
XX
XX 01-JUL-2002 (first entry)
XX
```

```
DE Chemically treated cell signalling DNA sequence complementary to#133.
XX
XX Cell signalling; cytosine methylation; cell signalling disease;
XX cancer; tumour; cytostatic; ds.
XX
XX Unidentified.
XX
XX WO200202807-A2.
XX
XX 10-JAN-2002.
XX
XX 29-JUN-2001: 2001WO-EP07471.
XX
XX 30-JUN-2000: 2000DE-1032529.
XX
XX 01-SEP-2000: 2000DE-1043826.
XX
XX (EPIC-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI: 2002-154758/20.
XX
XX Nucleic acid, useful for diagnosis and therapy of diseases associated
XX with cell signalling e.g. cancer, comprises chemically modified genomic
XX sequences of genes associated with cell signalling
XX
XX Claim 1: SEQ ID NO 266; 24pp+sequence listing; English.
XX
XX The invention relates to a nucleic acid comprising a sequence of at least
XX 18 bases of a segment of chemically pretreated DNA of genes associated
XX with cell signalling. The activity of the modified sequences of the
XX invention may be described as cytostatic. The object of the invention is
XX to provide the chemically modified DNA of genes associated with cell
XX signalling, as well as oligonucleotides and/or PNA-oligomers for
XX detecting cytosine methylations, as well as a method which is
XX particularly suitable for the diagnosis and/or therapy of genetic and
XX epigenetic parameters of genes associated with cell signalling. The
XX chemically modified DNA provided by the invention is useful for diagnosis
XX and therapy of diseases such as solid tumours and cancer. The sequences
XX given in records ABL70111-ABL70626 represent chemically pre-treated
XX genomic DNA's of genes associated with cell signalling.
XX Note: The sequence data for this patent is not represented in the printed
XX specification, but is based on sequence information supplied by the
XX European Patent Office.
XX
XX Sequence 16258 BP; 4242 A; 330 C; 3801 G; 7885 T; 0 other;
XX
XX Query Match 92.0%; Score 18.4; DB 24; Length 16258;
XX Best Local Similarity 95.0%; Pred. No. 68;
XX Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 TTAGTGTAGTTATTAGTT 20
XX ||||||||||||||||
XX DB 16029 TTAGTGTAGTTTATTAGTT 16048
XX
XX RESULT 12
XX ABL40038
XX ID ABL40038 standard; DNA: 16258 BP.
XX
XX ABL40038;
XX
XX 21-MAY-2002 (first entry)
XX
XX Human chemically pretreated gene sequence #60 strand 2.
XX
XX Human; ds; bisulphite treatment; CPG; DNA methylation; cancer; tumour;
XX cytostatic; ALDH6; CYP11A; CYP11B; CYP3A3; DPTD; EPH2; OCLN; TXNRD1;
XX UGT8; MRP; pharmacogenomics; SNP; single nucleotide polymorphism.
XX
XX Homo sapiens.
XX
XX WO200202806-A2.
XX
```

XX 10-JAN-2002.  
PD  
XX  
PE 29-JUN-2001; 2001WO-EP07470.  
XX  
PR 30-JUN-2000; 2000DE-1032529.  
XX  
PR 01-SEP-2000; 2000DE-1043826.  
XX  
PA (EPIC-) EPIGENOMICS AG.  
XX  
PI Olek A, Piepenbrock C, Berlin K;  
DR WPI: 2002-154757/20.  
XX  
XX New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,  
PT useful for detecting cytosine methylation state of genes associated  
PT with pharmacogenomics and for therapy of diseases e.g. cancer  
XX  
PS Claim 1: SEQ ID No 120; 24pp; English.  
XX  
CC The invention relates to a nucleic acid comprising a sequence at  
CC least 18 bases in length of a segment of the chemically pretreated DNA  
CC of genes associated with pharmacogenomics according to one of the  
CC sequences of the genes ALDH6 (NM\_000693), CYP11A (NM\_000781), CYP11B1  
CC (NM\_000497), CYP3A3 (NM\_000776 and NM\_017460), DYPD (NM\_000110), EPHX2  
CC (NM\_001979), OCLN (NM\_002538), TXNRD1 (NM\_003330), UGT8 (NM\_003360),  
CC MRP (NM\_004996, NM\_019900, NM\_019901, NM\_019902, NM\_019862, NM\_019898,  
CC NM\_019899) and their complementary sequences, or a sequence (SI) chosen  
CC from 87 sequences and their complementary sequences. The chemical pretreatment  
CC is bisulphite treatment to convert cytosines (but not methyl-cytosines)  
CC into uracils. Also included are an oligomer (II) in particular an  
CC oligonucleotide or a peptide nucleic acid (PNA)-oligomer, comprising in  
CC each case at least one base sequence having a length of 9 nucleotides  
CC which hybridises to or is identical to a chemically pretreated DNA of  
CC genes associated with pharmacogenomics and their complements, arranged in  
CC an array for analysing diseases associated with the polymorphisms)  
CC (CpG) and/or detecting SNPs (single nucleotide polymorphisms)  
CC of the 87 sequences. The oligomers may also be used as PCR primers.  
CC The set of 87 nucleic acids and their complements is useful for diagnosis  
CC and therapy of solid tumours and cancer. The present sequence  
CC represents one the 87 DNA sequences or its complement.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pcl\_sequences.  
XX  
SQ Sequence 16258 BP; 4242 A; 330 C; 3801 G; 7885 T; 0 other;  
Query Match 92.0%; Score 18.4; DB 24; Length 16258;  
Best Local Similarity 95.0%; Pred. No. 68;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TTAGTTGTTAGTTATTAGTT 20  
|||  
Db 16029 TTAGTTGTTAGTTATTAGTT 16048  
RESULT 13  
AAK65476/C  
ID AAK65476 standard; DNA; 27435 BP.  
XX  
AC AAK65476;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:20288.  
XX  
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200157182-A2.

XX 09-AUG-2001.  
PD  
XX  
PE 17-JAN-2001; 2001WO-US01354.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225578.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226682.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228824.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233603.  
PR 14-SEP-2000; 2000US-0233604.  
PR 14-SEP-2000; 2000US-0233605.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.

PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 01-NOV-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250381.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX

DR WPI; 2001-483426/52.  
XX  
XX Nucleic acids encoding human immune/haematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and  
PT metastasis -  
XX  
XX  
PS Disclosure; SEQ ID NO 20288; 3071pp + Sequence Listing; English.  
XX  
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patients own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention.  
XX  
SQ Sequence 27435 BP; 7241 A; 5321 C; 5251 G; 9622 T; 0 other:  
XX  
Query Match 90.0%; Score 18; DB 22; Length 27435;  
Best Local Similarity 100.0%; Pred. No. 96;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTAGTGTAGTATATAG 18  
DB 1648 TTACTGTAGTATATAG 1631  
IIIIIIIIIIIIIIIIIIII  
RESULT 14  
AAF28871  
ID AAF28871 standard; DNA; 20 BP.  
XX  
AC AAF28871;  
XX  
DT 09-MAY-2001 (first entry)  
XX  
XX Generic immunostimulatory oligonucleotide as vaccine adjuvant.  
DE  
XX Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;  
KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;  
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor; ss.  
XX  
OS Synthetic.  
XX  
PN WO200075304-A1.  
XX  
PD 14-DEC-2000.  
XX  
XX 08-JUN-2000; 2000WO-EP01566.  
PF  
XX 08-JUN-1999; 99FR-0007457.  
PR 06-AUG-1999; 99FR-0010378.  
XX  
XX (AVER ) AVENTIS PASTEUR.  
PA  
XX Bachy M, Sodayer R, Tranney E;  
PI  
XX WPI; 2001-041317/05.  
DR  
XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in  
PT vaccines for human use, induce lymphocyte proliferation and cytokine  
PT secretion -  
XX  
PS Claim 8; Page 17; 30pp; French.





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 11:12:37 ; Search time 45.25 Seconds  
(without alignments)  
135.548 Million cell updates/sec

Title: US-09-980-265-12

Perfect score: 20

Sequence: 1 ttatgtctgattgattatgtt 20

Scoring table: IDENTITY\_NUC

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued\_patents\_NA:\*

- 1: /cgn2\_6/ptodata/2/1na/5A\_COMB.seq:\*
- 2: /cgn2\_6/ptodata/2/1na/5B\_COMB.seq:\*
- 3: /cgn2\_6/ptodata/2/1na/6A\_COMB.seq:\*
- 4: /cgn2\_6/ptodata/2/1na/6B\_COMB.seq:\*
- 5: /cgn2\_6/ptodata/2/1na/PCrUS\_COMB.seq:\*
- 6: /cgn2\_6/ptodata/2/1na/backfillseq.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16.4	82.0	348	1	US-08-220-606B-54
2	15.8	79.0	5526	3	US-08-751-359-21
3	15.8	79.0	5526	4	US-08-907-146-21
4	15.8	79.0	17327	1	US-07-906-871-15
5	15.2	76.0	1578	4	US-09-134-001C-677
6	15.2	76.0	2509	1	US-08-047-413-14
7	15.2	76.0	2509	3	US-08-229-050-14
8	15.2	76.0	2509	3	US-08-801-563-14
9	15.2	76.0	2823	1	US-08-398-008A-1
10	15.2	76.0	2823	2	US-08-893-333-1
11	15.2	76.0	2862	1	US-08-148-209A-1
12	15.2	76.0	4483	1	US-08-181-271A-5
13	15.2	76.0	4483	1	US-08-449-315-5
14	15.2	76.0	4483	1	US-08-444-803-5
15	15.2	76.0	4483	1	US-08-449-043-5
16	15.2	76.0	4483	1	US-08-456-262A-5
17	15.2	76.0	4483	1	US-08-456-265A-5
18	15.2	76.0	4483	1	US-08-455-244-5
19	15.2	76.0	4483	1	US-08-454-876-5
20	15.2	76.0	4483	2	US-08-457-364-5
21	15.2	76.0	4483	2	US-08-456-262-5
22	15.2	76.0	4483	2	US-08-456-240-5
23	15.2	76.0	4483	2	US-08-455-736-5
24	15.2	76.0	4483	2	US-08-971-217-5
25	15.2	76.0	4483	4	US-09-350-600-5
26	15.2	76.0	4699	1	US-08-181-271A-6
27	15.2	76.0	4699	1	US-08-449-315-6

28	15.2	76.0	4699	1	US-08-444-803-6	Sequence 6, Appl1
29	15.2	76.0	4699	1	US-08-449-043-6	Sequence 6, Appl1
30	15.2	76.0	4699	1	US-08-456-265A-6	Sequence 6, Appl1
31	15.2	76.0	4699	1	US-08-455-416-6	Sequence 6, Appl1
32	15.2	76.0	4699	1	US-08-453-244-6	Sequence 6, Appl1
33	15.2	76.0	4699	1	US-08-454-876-6	Sequence 6, Appl1
34	15.2	76.0	4699	2	US-08-457-364-6	Sequence 6, Appl1
35	15.2	76.0	4699	2	US-08-456-262-6	Sequence 6, Appl1
36	15.2	76.0	4699	2	US-08-456-240-6	Sequence 6, Appl1
37	15.2	76.0	4699	2	US-08-455-736-6	Sequence 6, Appl1
38	15.2	76.0	4699	2	US-08-971-217-6	Sequence 6, Appl1
39	15.2	76.0	4699	4	US-09-350-600-6	Sequence 6, Appl1
40	15.2	76.0	59065	4	US-09-813-817-3	Sequence 3, Appl1
41	15.2	76.0	59065	4	US-09-978-197-3	Sequence 3, Appl1
42	14.8	74.0	200	2	US-08-875-972-27	Sequence 27, Appl1
43	14.8	74.0	203	4	US-09-605-785-717	Sequence 717, App
44	14.8	74.0	292	1	US-08-644-664B-41	Sequence 41, Appl1
45	14.8	74.0	292	2	US-08-761-277A-41	Sequence 41, Appl1

## ALIGNMENTS

RESULT 1  
US-08-220-606B-54  
Sequence 54, Application US/08220606B  
Patent No. 5641661  
GENERAL INFORMATION:  
APPLICANT: Kumagai, Monto H.  
APPLICANT: Genadi, Sverlow J.  
TITLE OF INVENTION: P1chia Pastoris Alcohol Oxidase Z2A1 and  
TITLE OF INVENTION: Z2A2 Regulatory Regions for Heterologous Gene Expression  
NUMBER OF SEQUENCES: 57  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/220,606B  
FILING DATE: 25-MAR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Halluin, Albert P.  
REGISTRATION NUMBER: 25,227  
REFERENCE/DOCKET NUMBER: 8129-065  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-854-3660  
TELEFAX: 415-854-3694  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 54:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 348 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA  
US-08-220-606B-54  
Query Match 82.0%; Score 16.4; DB 1; Length 348;  
Best Local Similarity 94.4%; Pred. No. 54;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 TAGTTGTTAGTTATTAGT 19  
DB 168 TAGTTGTTAGTTATTGTT 185

RESULT 2  
US-08-751-359-21  
Sequence 21, Application US/08751359  
Patent No. 6143559  
GENERAL INFORMATION:  
APPLICANT: Michael, Nancy M  
APPLICANT: Accavittl, Marianne  
APPLICANT: Thompson, Craig B  
TITLE OF INVENTION: METHODS FOR THE PRODUCTION OF CHICKEN  
TITLE OF INVENTION: MONOCLONAL ANTIBODIES  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESS: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/751,359  
FILING DATE: Concurrently Herewith  
CLASSIFICATION: 510  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: ARSB:504  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5526 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-751-359-21  
Query Match 79.0%; Score 15.8; DB 3; Length 5526;  
Best Local Similarity 89.5%; Pred. No. 92;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 TTAGTGTGTAGTTATTACT 19  
|||||  
Db 3120 TTAGTGTGTAGTTAGT 3138  
RESULT 3  
US-08-907-146-21  
Sequence 21, Application US/08907146  
Patent No. 6316600  
GENERAL INFORMATION:  
APPLICANT: Michael, Nancy M  
APPLICANT: Accavittl, Marianne  
APPLICANT: Thompson, Craig B  
TITLE OF INVENTION: METHODS FOR THE PRODUCTION OF CHICKEN  
TITLE OF INVENTION: MONOCLONAL ANTIBODIES  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESS: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/907,146  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/751,359  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: ARSB:504  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5526 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-907-146-21  
Query Match 79.0%; Score 15.8; DB 4; Length 5526;  
Best Local Similarity 89.5%; Pred. No. 92;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 TTAGTGTGTAGTTATTACT 19  
|||||  
Db 3120 TTAGTGTGTAGTTAGT 3138  
RESULT 4  
US-07-906-871-15  
Sequence 15, Application US/07906871  
Patent No. 5340739  
GENERAL INFORMATION:  
APPLICANT: Stevens, Richard L.  
APPLICANT: Abraham, Shalom  
TITLE OF INVENTION: HEMATOPOIETIC CELL SPECIFIC  
TITLE OF INVENTION: TRANSCRIPTIONAL REGULATORY ELEMENTS OF SRGLXCN AND USES  
TITLE OF INVENTION: THEREOF  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESS: Sterne, Kessler, Goldstein & Fox  
STREET: 1225 Connecticut Avenue, N.W., Suite 300  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/906,871  
FILING DATE: 19920103  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/816,289  
FILING DATE: 03 JAN 1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/635,544  
FILING DATE: 18-JAN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US89/03051  
FILING DATE: 13-JUL-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/224,035  
FILING DATE: 13-JUL-1988

ATTORNEY/AGENT INFORMATION:  
NAME: Cimbalá, Michele A  
REGISTRATION NUMBER: 33,851  
REFERENCE/DOCKET NUMBER: 0627,2830004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)833-7533  
TELEFAX: (202)833-8716  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17327 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: both  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
FEATURE:  
NAME/KEY: exon  
LOCATION: 621..753  
FEATURE:  
NAME/KEY: intron  
LOCATION: 754..9596  
FEATURE:  
NAME/KEY: exon  
LOCATION: 9597..9744  
FEATURE:  
NAME/KEY: intron  
LOCATION: 9745..16396  
FEATURE:  
NAME/KEY: exon  
LOCATION: 16397..17327  
US-07-906-871-15

Query Match 79.0%; Score 15.8; DB 1; Length 17327;  
Best Local Similarity 89.5%; Pred. No. 90;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAGTTGTTAGTTATTAGT 19  
DB 15437 TTAGTTGTTAGTTGTTGTTG 15455

RESULT 5  
US-09-134-001C-677  
Sequence 677, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
PRIOR FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 677  
LENGTH: 1578  
TYPE: DNA  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-677

Query Match 76.0%; Score 15.2; DB 4; Length 1578;  
Best Local Similarity 85.0%; Pred. No. 1,7e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTTGTTAGTTATTAGT 20  
DB 796 TTGTTATTAGTTATTGTT 815

RESULT 6  
US-08-047-413-14

Sequence 14, Application US/08047413  
Patent No. 5670706  
GENERAL INFORMATION:  
APPLICANT: Cornelissen, Bernardus J.C.  
APPLICANT: Melchers, Leo S.  
APPLICANT: Meulenhoff, Elisabeth J.S.  
APPLICANT: van Roekel, Jeroen S.C.  
APPLICANT: Sela-Buurlage, Marianne B.  
APPLICANT: Vloemans, Alexandra A.  
APPLICANT: Woloshuk, Charles P.  
APPLICANT: Bol, John F.  
TITLE OF INVENTION: FUNGAL RESISTANT PLANTS, PROCESS FOR  
TITLE OF INVENTION: OBTAINING FUNGAL RESISTANT PLANTS AND RECOMBINANT  
TITLE OF INVENTION: POLYNUCLEOTIDES FOR USE THEREIN  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 755 Page Mill Road  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/047,413  
FILING DATE: 19-APR-1993  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/647,831  
FILING DATE: 29-JAN-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Murashige, Kate H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 24615-20022,00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-813-5600  
TELEFAX: 415-494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2509 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(85..142, 930..1948)  
US-08-047-413-14

Query Match 76.0%; Score 15.2; DB 1; Length 2509;  
Best Local Similarity 85.0%; Pred. No. 1,6e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTTGTTAGTTATTAGT 20  
DB 215 TTAGTTGATTAGTTATTGCT 234

RESULT 7  
US-08-229-050-14  
Sequence 14, Application US/08229050  
Patent No. 6066491  
GENERAL INFORMATION:  
APPLICANT: Cornelissen, Bernardus J.C.  
APPLICANT: Melchers, Leo S.  
APPLICANT: Meulenhoff, Elisabeth J.S.  
APPLICANT: van Roekel, Jeroen S.C.  
APPLICANT: Sela-Buurlage, Marianne B.

```

? APPLICANT: Vloemans, Alexandra A.
? APPLICANT: Woloshuk, Charles P.
? APPLICANT: Bol, John F.
? APPLICANT: Linthorst, Hubertus J.M.
? TITLE OF INVENTION: FUNGAL RESISTANT PLANTS, PROCESS FOR
? TITLE OF INVENTION: OBTAINING FUNGAL RESISTANT PLANTS AND RECOMBINANT
? TITLE OF INVENTION: POLYNUCLEOTIDES FOR USE THEREIN
? NUMBER OF SEQUENCES: 17
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Morrison & Foerster
? STREET: 755 Page Mill Road
? CITY: Palo Alto
? STATE: California
? COUNTRY: USA
? ZIP: 94304-1018
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentln Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/229,050
? FILING DATE:
? CLASSIFICATION: 800
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/647,831
? FILING DATE: 29-JAN-1991
? ATTORNEY/AGENT INFORMATION:
? NAME: Murashige, Kate H.
? REGISTRATION NUMBER: 29,959
? REFERENCE/DOCKET NUMBER: 24615-20022.00
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 415-813-5600
? TELEFAX: 415-494-0792
? TELEX: 706141
? INFORMATION FOR SEQ. ID NO: 14:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2509 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? FEATURE:
? NAME/KEY: CDS
? LOCATION: join(85..142, 930..1948)
?
US-08-229-050-14
?
Query Match 76.0%; Score 15.2; DB 3; Length 2509;
Best Local Similarity 85.0%; Pred. No. 1.6e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTTGTTAGTTATTAGTT 20
Db 215 TTAGTTGATGATTATTGCT 234

RESULT 8
US-08-801-563-14
? Sequence 14, Application US/08801563
? Patent No. 6087560
? GENERAL INFORMATION:
? APPLICANT: Cornelissen, Bernardus J.C.
? APPLICANT: Meichers, Leo S.
? APPLICANT: Meulenhoff, Elisabeth J.S.
? APPLICANT: van Roekel, Jeroen S.C.
? APPLICANT: Sela-Buurlage, Marianne B.
? APPLICANT: Vloemans, Alexandra A.
? APPLICANT: Woloshuk, Charles P.
? APPLICANT: Bol, John F.
? APPLICANT: Linthorst, Hubertus J.M.
? TITLE OF INVENTION: FUNGAL RESISTANT PLANTS, PROCESS FOR
? TITLE OF INVENTION: OBTAINING FUNGAL RESISTANT PLANTS AND RECOMBINANT
? TITLE OF INVENTION: POLYNUCLEOTIDES FOR USE THEREIN
? NUMBER OF SEQUENCES: 17
```

```

? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Morrison & Foerster
? STREET: 755 Page Mill Road
? CITY: Palo Alto
? STATE: California
? COUNTRY: USA
? ZIP: 94304-1018
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentln Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/801,563
? FILING DATE:
? CLASSIFICATION: 800
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/047,413
? FILING DATE: 19-APR-1993
? APPLICATION NUMBER: US 07/647,831
? FILING DATE: 29-JAN-1991
? ATTORNEY/AGENT INFORMATION:
? NAME: Murashige, Kate H.
? REGISTRATION NUMBER: 29,959
? REFERENCE/DOCKET NUMBER: 24615-20022.00
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 415-813-5600
? TELEFAX: 415-494-0792
? TELEX: 706141
? INFORMATION FOR SEQ. ID NO: 14:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2509 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? FEATURE:
? NAME/KEY: CDS
? LOCATION: join(85..142, 930..1948)
?
US-08-801-563-14
?
Query Match 76.0%; Score 15.2; DB 3; Length 2509;
Best Local Similarity 85.0%; Pred. No. 1.6e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTTGTTAGTTATTAGTT 20
Db 215 TTAGTTGATGATTATTGCT 234

RESULT 9
US-08-398-008A-1
? Sequence 1, Application US/08398008A
? Patent No. 5665588
? GENERAL INFORMATION:
? APPLICANT: Kornbluth, Jack I
? TITLE OF INVENTION: DNA Encoding Natural Killer Lytic Associated
? TITLE OF INVENTION: Protein
? NUMBER OF SEQUENCES: 17
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Gildbreth & Adler, P.C.
? STREET: 8011 Candle Lane
? CITY: Houston
? STATE: Texas
? COUNTRY: USA
? ZIP: 77071
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
? OPERATING SYSTEM: Macintosh
? SOFTWARE: Microsoft Word 5.1
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/398,008A
? FILING DATE: March 2, 1995
```

CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/126,501  
FILING DATE: 24-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Adler, Dr. Benjamin Aaron  
REGISTRATION NUMBER: 35,423  
REFERENCE/DOCKET NUMBER: D5705CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (713) 777-2321  
TELEFAX: (713) 777-6908  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2823 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double-stranded  
TOPOLOGY: unknown  
MOLECULE TYPE: cDNA  
HYPOTHEICAL: no  
ANTI-SENSE: no  
US-08-398-008A-1

Query Match 76.0%; Score 15.2; DB 1; Length 2823;  
Best Local Similarity 85.0%; Pred. No. 1.6e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTTGTAGTTATTAGTT 20  
||||| | | | | | | | | |  
Db 2654 TTAGTTTAATTATTAGTT 2673

RESULT 10  
US-08-893-333-1  
Sequence 1, Application US/08893333A  
Patent No. 5981705  
GENERAL INFORMATION:  
APPLICANT: Kornbluth, Jacki  
TITLE OF INVENTION: DNA Encoding Natural Killer Lytic Associated Protein  
FILE REFERENCE: D5705CIP/D  
CURRENT APPLICATION NUMBER: US/08/893,333A  
CURRENT FILING DATE: 1997-07-16  
NUMBER OF SEQ ID NOS: 17  
SEQ ID NO 1  
LENGTH: 2823  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
LOCATION: 190..1953  
OTHER INFORMATION: CDS  
US-08-893-333-1

Query Match 76.0%; Score 15.2; DB 2; Length 2823;  
Best Local Similarity 85.0%; Pred. No. 1.6e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTTGTAGTTATTAGTT 20  
||||| | | | | | | | | |  
Db 2654 TTAGTTTAATTATTAGTT 2673

RESULT 11  
US-08-148-209A-1/c  
Sequence 1, Application US/08148209A  
Patent No. 5556780  
GENERAL INFORMATION:  
APPLICANT: Dzaou, Victor J  
APPLICANT: Makiyama, Masashi  
TITLE OF INVENTION: TYPE-2 ANGIOTENSIN II RECEPTOR AND GENE  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/148,209A  
FILING DATE: 05-NOV-1993  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Rowland, Bertram I  
REGISTRATION NUMBER: 20,015  
REFERENCE/DOCKET NUMBER: A-58491-1/BIR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2862-base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 132..1223  
US-08-148-209A-1

Query Match 76.0%; Score 15.2; DB 1; Length 2862;  
Best Local Similarity 85.0%; Pred. No. 1.6e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTTGTAGTTATTAGTT 20  
||||| | | | | | | | | |  
Db 1507 TTGTGTGAATTATATAGTT 1488

RESULT 12  
US-08-181-271A-5  
Sequence 5, Application US/08181271A  
Patent No. 5614395  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesing, John H.  
APPLICANT: Friedrich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Meins, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperison, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Uknes, Scott J.  
APPLICANT: Ward, Eric R.  
APPLICANT: Williams, Shericca C.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA

ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/181,271A  
FILING DATE: 13-JAN-94  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/PI/CGC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ. ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4483 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-181-271A-5

Query Match 76.0%; Score 15.2; DB 1; Length 4483;  
Best Local Similarity 85.0%; Pred. No. 1.6e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTTGTTAGTTATTGCTT 20

DB 1627 TTAGTTGATGATATTGCTT 1646  
||||| |||||||  
RESULT 13  
US-08-449-315-5  
Sequence 5, Application US/08449315  
Patent No. 5650505  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesting, John H.  
APPLICANT: Friedrich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Meins, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperison, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Uknes, Scott J.  
APPLICANT: Ward, Eric R.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/449,315  
FILING DATE: 24-MAY-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/181,271  
FILING DATE: 13-JAN-94  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506



FILED DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILED DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILED DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILED DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILED DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILED DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4483 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-449-315-5

Query Match 76.0%; Score 15.2; DB 1; Length 4483;  
Best Local Similarity 85.0%; Pred. No. 1.6e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 TTAGTTGTAGTTATTAGTT 20  
||||| |||||||  
Db 1627 TTAGTTGATGATTATTGCT 1646

RESULT 14  
US-08-444-803-5  
Sequence 5, Application US/08444803  
Patent No. 5654414  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesing, John H.  
APPLICANT: Friedrich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Weins, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Speirson, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Ukenes, Scott J.  
APPLICANT: Ward, Eric R.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
NUMBER OF INVENTIONS: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/444,803  
FILED DATE: 19-MAY-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/181,271  
FILED DATE: 13-JAN-94  
APPLICATION NUMBER: US 08/093,301  
FILED DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILED DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILED DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILED DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILED DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILED DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILED DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILED DATE: 20-OCT 1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILED DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILED DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILED DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILED DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILED DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4483 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-444-803-5

Query Match 76.0%; Score 15.2; DB 1; Length 4483;  
Best Local Similarity 85.0%; Pred. No. 1.6e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TTAGTGTAGTATTAGT 20  
||||| |||||  
Db 1627 TTAGTGTAGTATTAGT 1646

RESULT 15  
US-08-449-043-5  
Sequence 5, Application US/08449043  
Patent No. 5689044  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesing, John H.  
APPLICANT: Friedrich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Meins, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperison, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Uknes, Scott J.  
APPLICANT: Ward, Eric R.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/449,043  
FILING DATE: 24-MAY-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/181,271  
FILING DATE: 13-JAN-94  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/PI/CGC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ. ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4483 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-449-043-5

Query Match 76.0%; Score 15.2; DB 1; Length 4483;  
Best Local Similarity 85.0%; Pred. No. 1.6e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TTAGTGTAGTATTAGT 20  
||||| |||||  
Db 1627 TTAGTGTAGTATTAGT 1646

Search completed: April 6, 2003, 11:56:32  
Job time : 50.25 secs



;; TITLE OF INVENTION: Soybean Cyst Nematode Resistance  
;; FILE REFERENCE: 38-10(13810)B  
;; CURRENT APPLICATION NUMBER: US/09/754, 853A  
;; CURRENT FILING DATE: 2001-01-05  
;; PRIOR APPLICATION NUMBER: US 60/174, 880  
;; PRIOR FILING DATE: 2000-01-07  
;; NUMBER OF SEQ ID NOS: 1119  
;; SEQ ID NO 4  
;; LENGTH: 513509  
;; TYPE: DNA  
;; ORGANISM: Glycine max  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (111805)..(113968)..(114684)..(115204)  
;; FEATURE:  
;; NAME/KEY: unsure  
;; LOCATION: (1)..(513509)  
;; OTHER INFORMATION: unsure at all n locations  
;; OTHER INFORMATION: Clone ID: 318013\_region\_A3  
US-09-754-853A-4

Query Match 84.0%; Score 16.8; DB 9; Length 513509;  
Best Local Similarity 90.0%; Pred. No. 4.5e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TTAGTGTAGTTATTAGTT 20  
Db 133170 TTAGTGTAGTTATTAGTT 133151

RESULT 3  
US-09-834-975-775/C  
;; Sequence 775, Application US/09834975  
;; Patent No. US20020110815A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Lillie, James  
;; APPLICANT: Brown, Jeffrey  
;; APPLICANT: Bolt, Andrew  
;; APPLICANT: Van Hufel, Christophe  
;; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS  
;; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY  
;; TITLE OF INVENTION: OF HUMAN CANCERS  
;; FILE REFERENCE: MRI-016B  
;; CURRENT APPLICATION NUMBER: US/09/834, 975  
;; CURRENT FILING DATE: 2001-04-13  
;; PRIOR APPLICATION NUMBER: 60/197, 538  
;; PRIOR FILING DATE: 2000-04-14  
;; NUMBER OF SEQ ID NOS: 1046  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 775  
;; LENGTH: 3376  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; LOCATION: (1)...(3376)  
;; OTHER INFORMATION: n = A,T,C or G  
US-09-834-975-775

Query Match 82.0%; Score 16.4; DB 10; Length 3376;  
Best Local Similarity 94.4%; Pred. No. 4.8e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 AGTTGTAGTTATTAGTT 20  
Db 3261 AGTTGTAGTTATTAGTT 3244

RESULT 4  
US-08-781-986A-994/C  
;; Sequence 994, Application US/08781986A  
;; Publication No. US20030054436A1  
;; GENERAL INFORMATION:

;; APPLICANT: Charles Kunsch  
;; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
;; NUMBER OF SEQUENCES: 5255  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Human Genome Sciences, Inc.  
;; STREET: 9410 Key West Avenue  
;; CITY: Rockville  
;; STATE: Maryland  
;; COUNTRY: USA  
;; ZIP: 20850  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage  
;; COMPUTER: HP Vectra 486/33  
;; OPERATING SYSTEM: MSDOS version 6.2  
;; SOFTWARE: ASCII text  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/781, 986A  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Benson, Bob  
;; REGISTRATION NUMBER: 30,446  
;; REFERENCE/DOCKET NUMBER: PB248PP  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (301) 309-8504  
;; TELEFAX: (301) 309-8512  
;; INFORMATION FOR SEQ ID NO: 994:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 551 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: linear  
US-08-781-986A-994

Query Match 79.0%; Score 15.8; DB 7; Length 551;  
Best Local Similarity 89.5%; Pred. No. 7.1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TTAGTGTAGTTATTAGTT 19  
Db 476 TTAGTGTAGTTATTAGTT 458

RESULT 5  
US-09-920-300A-827/C  
;; Sequence 827, Application US/09920300A  
;; Patent No. US20020136728A1  
;; GENERAL INFORMATION:  
;; APPLICANT: King, Gordon E.  
;; APPLICANT: Meagher, Madeline Joy  
;; APPLICANT: Xu, Jiangchun  
;; APPLICANT: Secrist, Heather  
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
;; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
;; FILE REFERENCE: 210121.547  
;; CURRENT APPLICATION NUMBER: US/09/920, 300A  
;; CURRENT FILING DATE: 2001-07-31  
;; NUMBER OF SEQ ID NOS: 1789  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 827  
;; LENGTH: 605  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-920-300A-827

Query Match 79.0%; Score 15.8; DB 10; Length 605;  
Best Local Similarity 89.5%; Pred. No. 7.1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 TTAGTGTAGTTATTAGTT 20

Db 578 TAGTAGTATTATTCGTT 560

RESULT 6  
US-10-033-528-827/c  
Sequence 827, Application US/10033528  
Patent No. US20020131971A1  
GENERAL INFORMATION:  
APPLICANT: King, Gordon E.  
APPLICANT: Meagher, Madeleine Joy  
APPLICANT: Xu, Jiangchun  
APPLICANT: Secrist, Heather  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
FILE REFERENCE: 210121.547c1  
CURRENT APPLICATION NUMBER: US/10/033,528  
CURRENT FILING DATE: 2001-12-26  
NUMBER OF SEQ ID NOS: 1896  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 827  
LENGTH: 605  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-033-528-827

Query Match  
Best Local Similarity 89.5%; DB 12: Length 605;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TAGTTGTTAGTTATTAGTT 20  
Db 578 TAGTAGTATTATTCGTT 560

RESULT 7  
US-09-938-842A-3055  
Sequence 3055, Application US/09938842A  
Patent No. US20020160378A1  
GENERAL INFORMATION:  
APPLICANT: Harper, Jeff  
APPLICANT: Kieps, Joel  
APPLICANT: Wang, Xun  
APPLICANT: Zhu, Tong  
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
FILE REFERENCE: SRIPI300-3  
CURRENT APPLICATION NUMBER: US/09/938,842A  
CURRENT FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: US 60/227,866  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: US 60/264,647  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/300,111  
PRIOR FILING DATE: 2001-06-22  
NUMBER OF SEQ ID NOS: 5379  
SEQ ID NO 3055  
LENGTH: 657  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-938-842A-3055

Query Match  
Best Local Similarity 89.5%; DB 9: Length 657;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAGTTGTTAGTTATTAGT 19  
Db 82 TTAGTTATTACTATTAGT 100

RESULT 8  
US-09-887-576-491

Sequence 491, Application US/09887576  
Patent No. US20020144047A1  
GENERAL INFORMATION:  
APPLICANT: Budworth, P.  
APPLICANT: Brown, D.  
APPLICANT: Chang, H.  
APPLICANT: Zhu, T.  
APPLICANT: Han, B.  
APPLICANT: Wang, X.  
APPLICANT: Cooper, Bret  
TITLE OF INVENTION: Promoters for regulation of plant expression  
FILE REFERENCE: 1360.0010S1  
CURRENT APPLICATION NUMBER: US/09/887,576  
CURRENT FILING DATE: 2001-06-25  
PRIOR APPLICATION NUMBER: US 60/213,848  
PRIOR FILING DATE: 2000-06-23  
PRIOR APPLICATION NUMBER: US 60/214,087  
PRIOR FILING DATE: 2000-06-23  
PRIOR APPLICATION NUMBER: US 60/258,692  
PRIOR FILING DATE: 2000-12-29  
NUMBER OF SEQ ID NOS: 875  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 491  
LENGTH: 1305  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-887-576-491

Query Match  
Best Local Similarity 89.5%; DB 10: Length 1305;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAGTTGTTAGTTATTAGT 19  
Db 753 TTAGTTATTACTATTAGT 771

RESULT 9  
US-09-887-576-549  
Sequence 549, Application US/09887576  
Patent No. US20020144047A1  
GENERAL INFORMATION:  
APPLICANT: Budworth, P.  
APPLICANT: Brown, D.  
APPLICANT: Chang, H.  
APPLICANT: Zhu, T.  
APPLICANT: Han, B.  
APPLICANT: Wang, X.  
APPLICANT: Cooper, Bret  
TITLE OF INVENTION: Promoters for regulation of plant expression  
FILE REFERENCE: 1360.0010S1  
CURRENT APPLICATION NUMBER: US/09/887,576  
CURRENT FILING DATE: 2001-06-25  
PRIOR APPLICATION NUMBER: US 60/213,848  
PRIOR FILING DATE: 2000-06-23  
PRIOR APPLICATION NUMBER: US 60/214,087  
PRIOR FILING DATE: 2000-06-23  
PRIOR APPLICATION NUMBER: US 60/258,692  
PRIOR FILING DATE: 2000-12-29  
NUMBER OF SEQ ID NOS: 875  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 549  
LENGTH: 1305  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-887-576-549

Query Match  
Best Local Similarity 89.5%; DB 10: Length 1305;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAGTTGTTAGTTATTAGT 19  
Db 753 TTAGTTATTACTATTAGT 771

Db 753 TTAGTTACTTATTAGT 771

RESULT 10

US-09-887-576-97  
; Sequence 97, Application US/09887576  
; Patent No. US20020144047A1  
; GENERAL INFORMATION:  
; APPLICANT: Budworth, P.  
; APPLICANT: Brown, D.  
; APPLICANT: Chang, H.  
; APPLICANT: Zhu, T.  
; APPLICANT: Han, B.  
; APPLICANT: Wang, X.  
; APPLICANT: Cooper, Bret  
; TITLE OF INVENTION: Promoters for regulation of plant expression  
; FILE REFERENCE: 1360.001US1  
; CURRENT APPLICATION NUMBER: US/09/887,576  
; CURRENT FILING DATE: 2001-06-25  
; PRIOR APPLICATION NUMBER: US 60/213,848  
; PRIOR FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: US 60/214,087  
; PRIOR FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: US 60/258,692  
; PRIOR FILING DATE: 2000-12-29  
; NUMBER OF SEQ ID NOS: 875  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 97  
; LENGTH: 2000  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-887-576-97

Query Match

Best Local Similarity 79.0%; Score 15.8; DB 10; Length 2000;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAGTTGTTAGTTATTAGT 19  
Db 1426 TTAGTTACTTATTAGT 1444

RESULT 11

US-09-836-607-45/c  
; Sequence 45, Application US/09836607  
; Patent No. US20020098541A1  
; GENERAL INFORMATION:  
; APPLICANT: Ni, Jian  
; APPLICANT: Ruben, Steven M.  
; TITLE OF INVENTION: TNFR Related Gene 12  
; FILE REFERENCE: PF490P1  
; CURRENT APPLICATION NUMBER: US/09/836,607  
; CURRENT FILING DATE: 2001-04-17  
; PRIOR APPLICATION NUMBER: 60/158,388  
; PRIOR FILING DATE: 2000-04-19  
; PRIOR APPLICATION NUMBER: 09/421,112  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: 60/104,950  
; PRIOR FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO: 45  
; LENGTH: 35414  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-836-607-45

Query Match

Best Local Similarity 79.0%; Score 15.8; DB 10; Length 35414;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAGTTGTTAGTTATTAGT 19  
||| ||||| ||||| |||

Db 16222 TTATTTGTTAGTTATTGCT 16204

RESULT 12

US-09-998-598-2130/c  
; Sequence 2130, Application US/09998598  
; Patent No. US20020150922A1  
; GENERAL INFORMATION:  
; APPLICANT: Stolk, John A.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Cheneault, Ruth A.  
; APPLICANT: Meagher, Madelein Joy  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.561  
; CURRENT APPLICATION NUMBER: US/09/998,598  
; CURRENT FILING DATE: 2001-11-16  
; NUMBER OF SEQ ID NOS: 2606  
; SOFTWARE: Corlax Invention Disclosure Database  
; SEQ ID NO: 2130  
; LENGTH: 380  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-998-598-2130

Query Match

Best Local Similarity 77.0%; Score 15.4; DB 10; Length 380;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GTTGTTAGTTATTAGTT 20  
Db 93 GTTGTTAGTCATTAGTT 77

RESULT 13

US-09-864-761-2265  
; Sequence 2265, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO  
; FILE REFERENCE: Aecomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2265
; LENGTH: 467
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL03383.24
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 6.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.4
US-09-864-761-2265

```

```

Query Match          77.0%; Score 15.4; DB 10; Length 467;
Best Local Similarity 94.1%; Pred. No. 1e+03; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1;

```

```

QY 3 AGTGTGTTAGTTATTAGT 19
   |||||||
DB 205 AGCTGTAGTTATTAGT 221

```

```

RESULT 14
US-09-960-352-11874/c
; Sequence 11874, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 11874
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 51-LIB2809-009-Q1-E1-E4
US-09-960-352-11874

```

```

Query Match          76.0%; Score 15.2; DB 10; Length 323;
Best Local Similarity 85.0%; Pred. No. 1.2e+03; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 3;

```

```

QY 1 TTAGTTGTTAGTTATTAGT 20
   |||||||
DB 88 TTGTTGTTAGTTAGAGT 69

```

```

RESULT 15
US-09-964-824A-191/c
; Sequence 191, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Horrigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 191
; LENGTH: 441
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-824A-191

```

```

Query Match          76.0%; Score 15.2; DB 10; Length 441;
Best Local Similarity 85.0%; Pred. No. 1.2e+03; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 3;

```

```

QY 1 TTAGTTGTTAGTTATTAGT 20
   |||||||
DB 242 TTGTTGTTGTTAGTTATT 223

```

```

Search completed: April 6, 2003, 12:14:20
Job time : 167 secs

```

**THIS PAGE BLANK (USPTO)**